

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 19:53:11 ; Search time 133 Seconds
(without alignments)
254.377 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFQKSNLRHQT.....KQCGKAFGCPNLRHRGRTH 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*
- 9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	378	7 ADL33662	Adl33662 Zinc fing
2	340	76.9	107	9 AEB77956	Aeb77956 Zinc fing
3	340	76.9	111	8 ADP70949	Adp70949 Novel zin
4	334	75.6	111	8 ADP70959	Adp70959 Novel zin
5	330	74.7	111	8 ADP70948	Adp70948 Novel zin
6	328	74.2	111	8 ADP70930	Adp70930 Novel zin
7	328	74.2	111	8 ADP70950	Adp70950 Novel zin
8	328	74.2	113	8 ADP70936	Adp70936 Novel zin
9	328	74.2	125	8 ADO48443	Ado48443 Human VEG
10	327	74.0	213	5 AEG59963	Aeg59963 Human DIT
11	327	74.0	483	5 AAU10794	Aau10794 Human zif
12	327	74.0	555	4 AEG22742	Aeg22742 Novel hum
13	327	74.0	581	8 AEN99867	Adn99867 Novel hum
14	327	74.0	670	5 AAE25293	Aae25293 Novel nuc
15	327	74.0	673	7 ADM06183	Adm06183 Human pro
16	327	74.0	728	6 ABU11645	Abu11645 Human MOD
17	327	74.0	905	8 AEN99866	Adn99866 Novel hum
18	327	74.0	905	8 AEN99577	Adn99577 Novel hum
19	326	73.8	211	7 ADL33678	Adl33678 Zinc fing
20	325	73.5	107	9 AEB77950	Aeb77950 Zinc fing
21	325	73.5	184	4 AAE06005	Aae06005 Zinc fing
22	325	72.6	184	4 AAE06002	Aae06002 Zinc fing
23	320	72.4	111	8 ADP70952	Adp70952 Novel zin
24	319	72.2	592	7 ADM04984	Adm04984 Human pro

25	319	72.2	622	4 AAM78947	Aam78947 Human pro
26	319	72.2	631	4 AAM79931	Aam79931 Human pro
27	319	72.2	632	4 ABG18386	Abg18386 Novel hum
28	318	71.9	363	8 ADG74180	Adg74180 Artificia
29	318	71.9	365	8 ADG74182	Adg74182 Artificia
30	318	71.9	365	8 ADG74177	Adg74177 Artificia
31	317	71.7	111	8 ADP70938	Adp70938 Novel zin
32	316	71.5	99	5 AAE21128	Aae21128 Zinc fing
33	316	71.5	99	7 AAE38667	Aae38667 Zinc fing
34	316	71.5	113	8 ADP70928	Adp70928 Novel zin
35	316	71.5	394	6 ABU96726	Abu96726 Human nuc
36	315	71.3	107	7 ADL33484	Adl33484 Artificia
37	315	71.3	173	4 AAB94061	Aab94061 Human pro
38	315	71.3	209	7 ADL33670	Adl33670 Zinc fing
39	315	71.3	412	8 ADR09901	Adr09901 Human pro
40	315	71.3	434	6 ABU96677	Abu96677 Human nuc
41	315	71.3	450	4 AAO5464	Aao5464 Human pol
42	315	71.3	499	7 ADM04269	Adm04269 Human pro
43	315	71.3	530	7 ADB65280	Adb65280 Human pro
44	315	71.3	553	7 ADC31307	Adc31307 Human nov
45	315	71.3	563	8 ADR10447	Adr10447 Human pro

ALIGNMENTS

RESULT 1
ADL33662
ID ADL33662 standard; protein; 378 AA.
XX
AC ADL33662;
DT 20-MAY-2004 (first entry)
XX
DE Zinc finger domain-containing protein - Neurol-p65.
XX
KW modified cell; artificial transcription factor; stress resistance;
KW phenotypic trait alteration; zinc finger domain.
XX
OS Unidentified.
XX
PN WO2003048345-A1.
XX
PD 12-JUN-2003.
XX
PF 07-DEC-2002; 2002WO-KR002309.
XX
PR 07-DEC-2001; 2001US-0338441P.
PR 26-APR-2002; 2002US-0376053P.
PR 02-AUG-2002; 2002US-0400904P.
PR 05-AUG-2002; 2002US-0401089P.
(TOOL-) TOOLGEN INC.
XX
PI Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;
PI Jang Y;
DR WPI; 2003-513760/48.
DR N-PSDB; ADL33661.
XX
New modified cell comprising a heterologous nucleic acid encoding an
artificial transcription factor that confers stress resistance, useful
for altering a phenotypic trait of a cell or organism.
PS Disclosure; SEQ ID NO 202; 169pp; English.
XX
The invention comprises a modified cell containing a heterologous nucleic
acid encoding an artificial transcription factor that confers stress
resistance to the modified cell. The modified cell of the invention is
useful for altering a phenotypic trait of a cell or organism. The present
amino acid sequence represents a protein which contains zinc finger
domains.
XX

```
SQ Sequence 378 AA;
Query Match 100.0%; Score 442; DB 7; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.3e-42;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 60
Db 33 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 92
Qy 61 GKAFGCPNLRHRGRTH 77
Db 93 GKAFGCPNLRHRGRTH 109

RESULT 2
ID AEB77956 standard; protein; 107 AA.
AC AEB77956;
XX
XX
DT 22-SEP-2005 (first entry)
DE Zinc finger protein T7.
KW gene expression; protein engineering.
XX
XX Escherichia coli.
OS
XX WO2005061705-A1.
XX
XX 07-JUL-2005.
XX
XX 23-DEC-2004; 2004WO-KR003420.
XX
XX 23-DEC-2003; 2003US-0532362P.
XX
XX (TOOL-) TOOLGEN INC.
XX
XX Kim J, Park K, Jang Y;
XX WPI; 2005-479455/48.
XX
XX Regulating expression of a gene in a prokaryotic cell by expressing the
XX nucleic acid encoding the artificial polypeptide in the cell under
XX conditions where the polypeptide is produced, binds to the target DNA
XX site and regulates the gene.
XX
XX Example 3; SEQ ID NO 53; 117pp; English.
XX
XX The invention relates to a method of regulating expression of a gene in a
XX prokaryotic cell which comprises expressing the nucleic acid encoding the
XX artificial polypeptide in the cell under conditions in which the
XX artificial polypeptide is produced, binds to the target DNA site and
XX regulates the gene. The method is useful for regulating expression of a
XX gene in a prokaryotic cell. The present sequence represents the amino
XX acid sequence of a zinc finger protein.
XX
SQ Sequence 107 AA;
Query Match 76.9%; Score 340; DB 9; Length 107;
Best Local Similarity 74.0%; Pred. No. 3.2e-31;
Matches 57; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 60
Db 31 CDHCGKAFVSSNLRHQRTHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 90
Qy 61 GKAFGCPNLRHRGRTH 77
Db 91 GKAFGCPNLRHRGRTH 107

SQ Sequence 111 AA;
Query Match 76.9%; Score 340; DB 8; Length 111;
Best Local Similarity 74.0%; Pred. No. 3.4e-31;
Matches 57; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 60
Db 3 CKTCQKFSRSDHLKTKTRHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 62
Qy 61 GKAFGCPNLRHRGRTH 77
Db 63 GKAFGCPNLRHRGRTH 79
```

```
RESULT 3
ID ADP70949 standard; protein; 111 AA.
XX
XX ADP70949;
XX
XX 09-SEP-2004 (first entry)
XX
XX Novel zinc finger protein-related DNA-binding domain SeqID57.
XX
XX DNA binding domain; zinc finger domain; VEGF; VEGF-A;
XX transcription activation domain; transcription repression domain;
XX protein transduction domain; p65; VP16 activation domain; Kid;
XX KOX repression domain; TAT protein; VP22 protein;
XX Antennapedia homeodomain; cytosstatic; antiinflammatory; angiogenesis;
XX neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;
XX ds.
XX
XX Unidentified.
XX
XX WO2004053130-A1.
XX
XX 24-JUN-2004.
XX
XX 09-DEC-2003; 2003WO-KR002693.
XX
XX 09-DEC-2002; 2002US-0431892P.
XX
XX (TOOL-) TOOLGEN INC.
XX
XX Kim J, Shin H, Kwon H;
XX WPI; 2004-468871/44.
XX
XX New polypeptide comprising a DNA binding domain that includes a plurality
XX of zinc finger domains, and at least two of the zinc finger domains each
XX include respective zinc finger domain motifs, useful for modulating
XX angiogenesis.
XX
XX Example 7; SEQ ID NO 57; 176pp; English.
XX
XX This invention relates to a novel polypeptide comprising a DNA binding
XX domain that includes a plurality of zinc finger domains, where the DNA
XX binding domain can bind to a site in a VEGF gene, and at least two of the
XX zinc finger domains each include respective zinc finger domain motifs
XX listed in the specification. The VEGF gene is the human VEGF-A gene. The
XX polypeptide regulates the VEGF gene expression. The polypeptide further
XX comprises a transcription activation domain, a transcription repression
XX domain, or a protein transduction domain. The transcription activation
XX domain comprises p65 or VP16 activation domain. The transcription
XX repression domain comprises Kid or KOX repression domain. The protein
XX transduction domain is a part of TAT protein, VP22 protein, or
XX Antennapedia homeodomain. The invention may be useful for the production
XX of compounds with a cytosstatic or antiinflammatory activity acting as
XX angiogenesis modulators. The polypeptides and nucleic acids are useful
XX for modulating angiogenesis whilst the compounds may be useful for
XX preventing or treating a neoplastic disorder, an inflammatory disorder,
XX or an angiogenesis-based disorder. The present sequence is that of a DNA-
XX binding domain which is related to the polypeptide of the invention.
XX
XX
SQ Sequence 111 AA;
Query Match 76.9%; Score 340; DB 8; Length 111;
Best Local Similarity 74.0%; Pred. No. 3.4e-31;
Matches 57; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 60
Db 3 CKTCQKFSRSDHLKTKTRHTGKPYKCECGKRAFTOSSNLTKHKIKHTGKPYKCKQC 62
Qy 61 GKAFGCPNLRHRGRTH 77
Db 63 GKAFGCPNLRHRGRTH 79
```

DB	91	GRAFNRRSHLTRHRIH 107	
RESULT 4			
ID	ADP70959	standard; protein; 111 AA.	
XX	AC	ADP70959;	
XX	DT	09-SEP-2004 (first entry)	
XX	DE	Novel zinc finger protein-related DNA-binding domain SeqID67.	
XX	DE	DNA binding domain; zinc finger domain; VEGF; VEGF-A;	
KW	KW	transcription activation domain; transcription repression domain;	
KW	KW	protein transduction domain; p65; VP16 activation domain; Kid;	
KW	KW	KOX repression domain; TAT protein; VP22 protein;	
KW	KW	Antennapedia homeodomain; cytosstatic; antiinflammatory; angiogenesis;	
KW	KW	neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;	
ds.			
XX	OS	Unidentified.	
XX	PN	WO2004053130-A1.	
XX	PD	24-JUN-2004.	
XX	PF	09-DEC-2003; 2003WO-KR002693.	
XX	PR	09-DEC-2002; 2002US-0431892P.	
XX	PA	(TOOL-) TOOLGEN INC.	
XX	PI	Kim J, Shin H, Kwon H;	
XX	PI	WPI; 2004-468871/44.	
XX	DR		
XX	PT	New polypeptide comprising a DNA binding domain that includes a plurality	
PT	PT	of zinc finger domains, and at least two of the zinc finger domains each	
PT	PT	include respective zinc finger domain motifs, useful for modulating	
XX	XX	angiogenesis.	
PS	PS	Example 7; SEQ ID NO 67; 176pp; English.	
XX	CC	This invention relates to a novel polypeptide comprising a DNA binding	
CC	CC	domain that includes a plurality of zinc finger domains, where the DNA	
CC	CC	binding domain can bind to a site in a VEGF gene, and at least two of the	
CC	CC	zinc finger domains each include respective zinc finger domain motifs	
CC	CC	listed in the specification. The VEGF gene is the human VEGF-A gene. The	
CC	CC	polypeptide regulates the VEGF gene expression. The polypeptide further	
CC	CC	comprises a transcription activation domain, a transcription repression	
CC	CC	domain, or a protein transduction domain. The transcription activation	
CC	CC	domain comprises p65 or VP16 activation domain. The transcription	
CC	CC	repression domain comprises Kid or KOX repression domain. The protein	
CC	CC	transduction domain is a part of TAT protein, VP22 protein, or	
CC	CC	Antennapedia homeodomain. The invention may be useful for the production	
CC	CC	of compounds with a cytostatic or antiinflammatory activity acting as	
CC	CC	angiogenesis modulators. The polypeptides and nucleic acids are useful	
CC	CC	for modulating angiogenesis whilst the compounds may be useful for	
CC	CC	preventing or treating a neoplastic disorder, an inflammatory disorder,	
CC	CC	or an angiogenesis-based disorder. The present sequence is that of a DNA-	
CC	CC	binding domain which is related to the polypeptide of the invention.	
XX	XX	Sequence 111 AA;	
Query Match	75.6%;	Score 334; DB 8; Length 111;	
Best Local Similarity	75.3%;	Pred. No. 1.6e-30;	
Matches	58; Conservative	7; Mismatches 12; Indels 0; Gaps 0;	
QY	1	CKDCGKAFIQKSNLIRHORTHGKPYKCECGKAFQSSNLTQKKLHTGKPYKCKQC 60	
DB	31	CPDCGKSFQSSLLIRHORTHGKPYKCECGKAFQSSNLTQKKLHTGKPYKCKQC 90	
QY	61	GRAFGCPNLRHGRTH 77	
DB	91	GRAFNRRSHLTRHRIH 107	
RESULT 5			
ID	ADP70948	standard; protein; 111 AA.	
XX	AC	ADP70948;	
XX	DT	09-SEP-2004 (first entry)	
XX	DE	Novel zinc finger protein-related DNA-binding domain SeqID56.	
XX	DE	DNA binding domain; zinc finger domain; VEGF; VEGF-A;	
KW	KW	transcription activation domain; transcription repression domain;	
KW	KW	protein transduction domain; p65; VP16 activation domain; Kid;	
KW	KW	KOX repression domain; TAT protein; VP22 protein;	
KW	KW	Antennapedia homeodomain; cytosstatic; antiinflammatory; angiogenesis;	
KW	KW	neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;	
ds.			
XX	OS	Unidentified.	
XX	PN	WO2004053130-A1.	
XX	PD	24-JUN-2004.	
XX	PF	09-DEC-2003; 2003WO-KR002693.	
XX	PR	09-DEC-2002; 2002US-0431892P.	
XX	PA	(TOOL-) TOOLGEN INC.	
XX	PI	Kim J, Shin H, Kwon H;	
XX	PI	WPI; 2004-468871/44.	
XX	DR		
XX	PT	New polypeptide comprising a DNA binding domain that includes a plurality	
PT	PT	of zinc finger domains, and at least two of the zinc finger domains each	
PT	PT	include respective zinc finger domain motifs, useful for modulating	
XX	XX	angiogenesis.	
PS	PS	Example 7; SEQ ID NO 56; 176pp; English.	
XX	CC	This invention relates to a novel polypeptide comprising a DNA binding	
CC	CC	domain that includes a plurality of zinc finger domains, where the DNA	
CC	CC	binding domain can bind to a site in a VEGF gene, and at least two of the	
CC	CC	zinc finger domains each include respective zinc finger domain motifs	
CC	CC	listed in the specification. The VEGF gene is the human VEGF-A gene. The	
CC	CC	polypeptide regulates the VEGF gene expression. The polypeptide further	
CC	CC	comprises a transcription activation domain, a transcription repression	
CC	CC	domain, or a protein transduction domain. The transcription activation	
CC	CC	domain comprises p65 or VP16 activation domain. The transcription	
CC	CC	repression domain comprises Kid or KOX repression domain. The protein	
CC	CC	transduction domain is a part of TAT protein, VP22 protein, or	
CC	CC	Antennapedia homeodomain. The invention may be useful for the production	
CC	CC	of compounds with a cytostatic or antiinflammatory activity acting as	
CC	CC	angiogenesis modulators. The polypeptides and nucleic acids are useful	
CC	CC	for modulating angiogenesis whilst the compounds may be useful for	
CC	CC	preventing or treating a neoplastic disorder, an inflammatory disorder,	
CC	CC	or an angiogenesis-based disorder. The present sequence is that of a DNA-	
CC	CC	binding domain which is related to the polypeptide of the invention.	
XX	XX	Sequence 111 AA;	
Query Match	74.7%;	Score 330; DB 8; Length 111;	
Best Local Similarity	71.4%;	Pred. No. 4.8e-30;	
Matches	55; Conservative	11; Mismatches 11; Indels 0; Gaps 0;	
QY	1	CKDCGKAFIQKSNLIRHORTHGKPYKCECGKAFQSSNLTQKKLHTGKPYKCKQC 60	
DB	31	CRKCGRGSFQSSLLIRHORTHGKPYKCECGKAFQSSNLTQKKLHTGKPYKCKQC 90	
QY	31	CRKCGRGSFQSSLLIRHORTHGKPYKCECGKAFQSSNLTQKKLHTGKPYKCKQC 90	

```

QY      61 GKAFGCPNLRHGRTH 77
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      91 GKAFWPSNLRHGRTH 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
ID ADP70930 standard; protein; 111 AA.
XX
AC ADP70930;
XX
DT 09-SEP-2004 (first entry)
XX
DE Novel zinc finger protein-related DNA-binding domain SeqID38.
XX
KW DNA binding domain; zinc finger domain; VEGF; VEGF-A;
KW transcription activation domain; transcription repression domain;
KW protein transduction domain; p65; VP16 activation domain; Kid;
KW KOX repression domain; TAR protein; VP22 protein;
KW Antennapedia homeodomain; cytosolic; antiinflammatory; angiogenesis;
KW neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;
KW ds.
XX
OS Unidentified.
XX
PN WO2004053130-A1.
XX
PD 24-JUN-2004.
XX
PF 09-DEC-2003; 2003WO-KR002693.
XX
PR 09-DEC-2002; 2002US-0431892P.
XX
PA (TOOL-) TOOLGEN INC.
XX
PI Kim J, Shin H, Kwon H;
XX
DR WPI; 2004-468871/44.
XX
PT New polypeptide comprising a DNA binding domain that includes a plurality
PT of zinc finger domains, and at least two of the zinc finger domains each
PT include respective zinc finger domain motifs, useful for modulating
PT angiogenesis.
PS Example 7; SEQ ID NO 38; 176pp; English.
XX
CC This invention relates to a novel polypeptide comprising a DNA binding
CC domain that includes a plurality of zinc finger domains, where the DNA
CC binding domain can bind to a site in a VEGF gene, and at least two of the
CC zinc finger domains each include respective zinc finger domain motifs
CC listed in the specification. The VEGF gene is the human VEGF-A gene. The
CC polypeptide regulates the VEGF gene expression. The polypeptide further
CC comprises a transcription activation domain, a transcription repression
CC domain, or a protein transduction domain. The transcription activation
CC domain comprises p65 or VP16 activation domain. The transcription
CC repression domain comprises Kid or KOX repression domain. The protein
CC transduction domain is a part of TAR protein, VP22 protein, or
CC Antennapedia homeodomain. The invention may be useful for the production
CC of compounds with a cytosolic or antiinflammatory activity acting as
CC angiogenesis modulators. The polypeptides and nucleic acids are useful
CC for modulating angiogenesis whilst the compounds may be useful for
CC preventing or treating a neoplastic disorder, an inflammatory disorder,
CC or an angiogenesis-based disorder. The present sequence is that of a DNA-
CC binding domain which is related to the polypeptide of the invention.
XX
SQ Sequence 111 AA;
Query Match 74.2%; Score 328; DB 8; Length 111;
Best Local Similarity 71.4%; Pred. No. 8.1e-30;
Matches 55; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

QY      1 CKDCGKAFIQSNLRHGRTH 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      91 GKAFWPSNLRHGRTH 107
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
ID ADP70950 standard; protein; 111 AA.
XX
AC ADP70950;
XX
DT 09-SEP-2004 (first entry)
XX
DE Novel zinc finger protein-related DNA-binding domain SeqID58.
XX
KW DNA binding domain; zinc finger domain; VEGF; VEGF-A;
KW transcription activation domain; transcription repression domain;
KW protein transduction domain; p65; VP16 activation domain; Kid;
KW KOX repression domain; TAR protein; VP22 protein;
KW Antennapedia homeodomain; cytosolic; antiinflammatory; angiogenesis;
KW neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;
KW ds.
XX
OS Unidentified.
XX
PN WO2004053130-A1.
XX
PD 24-JUN-2004.
XX
PF 09-DEC-2003; 2003WO-KR002693.
XX
PR 09-DEC-2002; 2002US-0431892P.
XX
PA (TOOL-) TOOLGEN INC.
XX
PI Kim J, Shin H, Kwon H;
XX
DR WPI; 2004-468871/44.
XX
PT New polypeptide comprising a DNA binding domain that includes a plurality
PT of zinc finger domains, and at least two of the zinc finger domains each
PT include respective zinc finger domain motifs, useful for modulating
PT angiogenesis.
PS Example 7; SEQ ID NO 58; 176pp; English.
XX
CC This invention relates to a novel polypeptide comprising a DNA binding
CC domain that includes a plurality of zinc finger domains, where the DNA
CC binding domain can bind to a site in a VEGF gene, and at least two of the
CC zinc finger domains each include respective zinc finger domain motifs
CC listed in the specification. The VEGF gene is the human VEGF-A gene. The
CC polypeptide regulates the VEGF gene expression. The polypeptide further
CC comprises a transcription activation domain, a transcription repression
CC domain, or a protein transduction domain. The transcription activation
CC domain comprises p65 or VP16 activation domain. The transcription
CC repression domain comprises Kid or KOX repression domain. The protein
CC transduction domain is a part of TAR protein, VP22 protein, or
CC Antennapedia homeodomain. The invention may be useful for the production
CC of compounds with a cytosolic or antiinflammatory activity acting as
CC angiogenesis modulators. The polypeptides and nucleic acids are useful
CC for modulating angiogenesis whilst the compounds may be useful for
CC preventing or treating a neoplastic disorder, an inflammatory disorder,
CC or an angiogenesis-based disorder. The present sequence is that of a DNA-
CC binding domain which is related to the polypeptide of the invention.
XX
SQ Sequence 111 AA;
Query Match 74.2%; Score 328; DB 8; Length 111;
Best Local Similarity 55.2%; Pred. No. 8.1e-30;
Matches 58; Conservative 10; Mismatches 9; Indels 28; Gaps 1;

```

QY 1 CNDGKAFIQKSNLIRHQRTHGKPKYKCECGKA----- 35
 Db 3 CMECGKAFNRSLRTRHQRTHGKPKYKCKQCGKAFGCPNLRHGRTHGKPFQCKTC 62
 QY 36 ---FTOSSNLTUKKIHGTGKPKYKCKQCGKAFGCPNLRHGRTH 77
 Db 63 QRKFSRDLKTRTHGKPKYKCKQCGKAFGCPNLRHGRTH 107

RESULT 8
 ADP70936 standard; protein; 113 AA.
 AC ADP70936;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Novel zinc finger protein-related DNA-binding domain SeqID44.
 XX
 KW DNA binding domain; zinc finger domain; VEGF; VEGF-A;
 KW transcription activation domain; transcription repression domain;
 KW protein transduction domain; p65; VP16 activation domain; Kid;
 KW KRX repression domain; RAR protein; VP22 protein;
 KW Antennapedia homeodomain; cytosolic; antiinflammatory; angiogenesis;
 KW neoplastic disorder; inflammatory disorder; angiogenesis-based disorder;
 KW ds.
 XX
 OS Unidentified.
 XX
 PN WO2004053130-A1.
 XX
 PD 24-JUN-2004.
 XX
 PF 09-DEC-2003; 2003WO-KR002693.
 XX
 PR 09-DEC-2002; 2002US-0431892P.
 XX
 PA (TOOL-) TOOLGEN INC.
 XX
 PI Kim J, Shin H, Kwon H;
 XX
 DR WPI; 2004-468871/44.
 XX
 PT New polypeptide comprising a DNA binding domain that includes a plurality
 of zinc finger domains, and at least two of the zinc finger domains each
 include respective zinc finger domain motifs, useful for modulating
 angiogenesis.
 XX
 PS Example 7; SEQ ID NO 44; 176pp; English.
 XX
 CC This invention relates to a novel polypeptide comprising a DNA binding
 domain that includes a plurality of zinc finger domains, where the DNA
 binding domain can bind to a site in a VEGF gene, and at least two of the
 zinc finger domains each include respective zinc finger domain motifs
 listed in the specification. The VEGF gene is the human VEGF-A gene. The
 polypeptide regulates the VEGF gene expression. The polypeptide further
 comprises a transcription activation domain, a transcription repression
 domain, or a protein transduction domain. The transcription repression
 domain comprises p65 or VP16 activation domain. The transcription
 repression domain comprises Kid or KRX repression domain. The protein
 transduction domain is a part of TAT protein, VP22 protein, or
 Antennapedia homeodomain. The invention may be useful for the production
 of compounds with a cytostatic or antiinflammatory activity acting as
 angiogenesis modulators. The polypeptides and nucleic acids are useful
 for modulating angiogenesis whilst the compounds may be useful for
 preventing or treating a neoplastic disorder, an inflammatory disorder,
 or an angiogenesis-based disorder. The present sequence is that of a DNA-
 binding domain which is related to the polypeptide of the invention.
 XX
 SQ Sequence 113 AA;

Best Local Similarity 75.3%; Pred. No. 8.3e-30;
 Matches 58; Conservative 6; Mismatches 13; Indels 0; Gaps 0;
 QY 1 CNDGKAFIQKSNLIRHQRTHGKPKYKCECGKAFGCPNLRHGRTHGKPFQCKTC 60
 Db 33 CPDCKGKSFQSSSLIRHQRTHGKPKYKCECGKAFGCPNLRHGRTHGKPFQCKTC 92
 QY 61 GRAFGCPNLRHGRTH 77
 Db 93 GRGFSRKSNIIRHQRTH 109

RESULT 9
 ADO48443
 ID ADO48443 standard; protein; 125 AA.
 XX ADO48443;
 AC
 DT 12-AUG-2004 (first entry)
 XX
 DE Human VEGF-A gene targeting zinc finger protein, ZFP-F109.
 XX
 KW regulating; target; zinc finger protein; VEGF-A; cytostatic;
 KW transcription; post-transcription.
 XX
 OS Unidentified.
 XX
 PN WO2004044202-A1.
 XX
 PD 27-MAY-2004.
 XX
 PF 14-NOV-2003; 2003WO-KR002451.
 XX
 PR 14-NOV-2002; 2002KR-00070845.
 XX
 PA (TOOL-) TOOLGEN INC.
 XX
 PI Kim J, Shin H, Kwon H;
 XX
 DR WPI; 2004-411728/38.
 XX
 PT Regulating a target gene at both transcription and post-transcription
 levels comprises introducing into a cell a zinc finger protein that can
 bind to a target gene or a nucleic acid encoding the protein, and a RNA
 molecule.
 XX
 PS Claim 5; SEQ ID NO 5; 75pp; English.
 XX
 CC The invention relates to a novel method for regulating a target gene. The
 invention further comprises: a composition for regulating a target gene,
 the composition comprising a zinc finger protein that can bind to the
 target gene or a nucleic acid encoding the protein; and a RNA molecule
 that includes a strand including a sequence complementary to an mRNA
 transcribed from the target gene; treating a neoplastic disorder in a
 subject comprising providing the composition, and administering the
 composition to the subject to modulate VEGF-A gene expression in a cell
 of the subject; a kit that comprises the zinc finger protein, and the RNA
 molecule; a modified eukaryotic cell that comprises the zinc finger
 protein, and the RNA molecule; and modulating expression of a target gene
 in a cell of a subject by administering a first component that comprises
 a zinc finger protein or a nucleic acid encoding the zinc finger protein,
 and a second component that comprises a RNA molecule or a nucleic acid
 that can produce the RNA molecule, to the subject, where the zinc finger
 protein can bind to the target gene and modulate transcription of the
 target gene, the RNA molecule can reduce translatability of a transcript
 of the target gene, and the first and second components are administered
 in amounts effective to modulate expression of the target gene in a cell
 of the subject. The method of the invention has cytostatic activity. The
 method is useful for regulating a target gene at both transcription and
 post-transcription levels. This sequence represents a human VEGF-A gene
 targeting zinc finger protein of the invention.
 XX
 SQ Sequence 125 AA;

Query Match 74.2%; Score 328; DB 8; Length 125;
Best Local Similarity 75.3%; Pred. No. 9.3e-30; Indels 0; Gaps 0;
Matches 58; Conservative 6; Mismatches 13; Mismatches 13; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSLNLRHQRTHTGKPYKCEGKAFQSSNLTGKHKIHTGKPKYCKQC 60
DB 42 CPDCGKSFQSSSLNLRHQRTHTGKPYKCEGKAFQSSNLTGKHKIHTGKPKYCKQC 101
QY 61 GKAFGCPNLRHGRTH 77
DB 102 GRGFSRKSINLRHQRTH 118

RESULT 10
ABG59963
ID ABG59963 standard; protein; 213 AA.
XX AC
AC ABG59963;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human DITHP polypeptide #21.
XX
KW Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
KW inflammatory disorder; viral infection; bacterial infection; seizure;
KW fungal infection; parasitic infections; developmental disorder; breast;
KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
KW gastrointestinal disorder; transport disorder; gene therapy; kidney;
KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
KW thymus.
XX
OS Homo sapiens.
XX
PN WO200220754-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US027127.
XX
PR 05-SEP-2000; 2000US-0229747P.
PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.
PR 05-SEP-2000; 2000US-0230583P.
PR 06-SEP-2000; 2000US-0230514P.
PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.
PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.
PR 06-SEP-2000; 2000US-0230595P.
PR 06-SEP-2000; 2000US-0230597P.
PR 06-SEP-2000; 2000US-0230598P.
PR 06-SEP-2000; 2000US-0230599P.
PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230865P.
PR 07-SEP-2000; 2000US-0230989P.
PR 07-SEP-2000; 2000US-0230951P.
PR 07-SEP-2000; 2000US-0231163P.
PR 07-SEP-2000; 2000US-0231167P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiyama MG, Bradley DB, Kohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX WPI; 2002-383054/41.

DR N-PSDB; ABK71555.
XX
PI An isolated polynucleotide useful in diagnostics and therapeutics.
XX
PS Claim 29; Page 543; 686pp; English.
XX
CC The invention relates to human diagnostic and therapeutic (dithp)
CC polynucleotides and their associated polypeptides (DITHP polypeptides).
CC The sequences of the invention are used in the treatment and diagnosis of
CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal
CC infections, parasitic infections, developmental disorders (e.g. anaemia,
CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
CC (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences
CC ABG59943-ABG50220 represent human DITHP polypeptides of the invention
XX
SQ Sequence 213 AA;
Query Match 74.0%; Score 327; DB 5; Length 213;
Best Local Similarity 72.7%; Pred. No. 2.2e-29;
Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSLNLRHQRTHTGKPYKCEGKAFQSSNLTGKHKIHTGKPKYCKQC 60
DB 104 CKQCGKAFGNSASHLMHGRTHTGKPYCKQCGSFGCASRLQMHGRTHTGKPKYCKQC 163
QY 61 GKAFGCPNLRHGRTH 77
DB 164 GKAFGCPNLRHGRTH 180

RESULT 11
AAU10794
ID AAU10794 standard; protein; 483 AA.
XX
AC AAU10794;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human zinc finger protein 53.
XX
KW Human; zinc finger protein 53; malignant tumour; HIV infection;
KW immunological disease; inflammation.
XX
OS Homo sapiens.
XX
PN CN1307009-A.
XX
PD 08-AUG-2001.
XX
PF 21-JAN-2000; 2000CN-00111472.
XX
PR 21-JAN-2000; 2000CN-00111472.
XX
PA (BODA-) BODAO GENE TECH CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-049871/07.
XX N-PSDB; AAS18786.
XX Polypeptide-human zinc finger protein 53 and polynucleotide for coding
PT said polypeptide.
XX
PS Claim 1; Page 26-27 (Disclosure); 37pp; Chinese.

PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 05-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471338P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486896P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kochakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
DR WPI; 2004-365511/34.
DR N-PSDB; ADN99083.
XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
PS Claim 14; SEQ ID NO 1467; 532pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
XX protein of the invention.
XX
SQ Sequence 581 AA;
Query Match 74.0%; Score 327; DB 8; Length 581;
Best Local Similarity 72.7%; Pred. No. 6.7e-29;
Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 1 CKDCGKAFIQSNLIRHQHRTHTGKPYKCECGKAPTQSSNLTGHKHTGKPYKCKQC 60

Db 472 CKQCGKAFGSGASHLQMHGRHTHTGKPYECKQSGFCASRLQMHGRHTGKPYKCKQC 531
QY 61 GKAFGCPNSLRRHGRTH 77
Db 532 GKAFGCPNSLRRHGRTH 548
RESULT 14
AAE25293
ID AAE25293 standard; protein; 670 AA.
XX
AC AAE25293;
XX
DT 30-OCT-2002 (first entry)
XX
DE Human nucleic acid-associated protein (NAAP-12).
XX
KW Human; nucleic acid-associated protein; NAAP-12; neurological disorder;
KW arteriosclerosis; cancer; cell proliferative disorder; atherosclerosis;
KW lymphoma; epilepsy; Alzheimer's disease; developmental; anticonvulsant;
KW autoimmune disorder; AIDS; allergy; anaemia; stroke; malaria; leishmania;
KW gene therapy; nootropic; neuroprotective; cerebroprotective; virucide;
KW immunosuppressive; protozoacide; antimicrobial.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1. 75 /note= "KRAB BOX domain"
FT Region 4. 74 /note= "KRAB BOX"
FT
XX WO200250279-A2.
XX
PD 27-JUN-2002.
XX
XX 19-DEC-2001; 2001WO-US050256.
XX
PR 21-DEC-2000; 2000US-0257714P.
PR 05-JAN-2001; 2001US-0260081P.
PR 16-JAN-2001; 2001US-0262302P.
PR 23-FEB-2001; 2001US-0263823P.
PR 02-FEB-2001; 2001US-0266088P.
PR 29-OCT-2001; 2001US-0348442P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Baughin MR, Lu Y, Arvizu C, Ramkumar J, Yao MG, Policky JL;
PI Walia NK, Tribouley KM, Yue H, Batra S, Ding L, Lal PG;
PI Borowsky ML, Lu DM, Gandhi AR, Griffin JA, Xu Y, Azimzai Y;
PI Gietzen KU, Tang YT, Warren BA, Mason PM, Burford N, Hafalia AJA;
PI Lee EA, Yang J, Gorvad AE, Emerling BM, Marquis JP, Lee SY;
PI Swarnakar A, Reddy R;
XX
DR WPI; 2002-519887/55.
DR N-PSDB; AAD41202.
XX
XX Nucleic acid associated proteins and nucleic acids for diagnosing,
PT treating and preventing cell proliferative (e.g. cancers), neurological
PT (e.g. epilepsy or stroke) or autoimmune disorders (e.g. AIDS).
XX
PS Claim 67; Page 167-169; 193pp; English.
XX
CC The invention relates to nucleic acid-associated proteins (NAAP) and
CC nucleic acids. The nucleic acid and amino acid sequences are useful for
CC diagnosing, treating and preventing cell proliferative e.g.
CC arteriosclerosis, atherosclerosis, lymphoma or cancer), neurological
CC (e.g. epilepsy, Alzheimer's disease or stroke), developmental, and
CC autoimmune disorders (e.g. AIDS, allergies, or anaemia) or infections
CC (e.g. malaria, or leishmania), as well as in assessing the effects of
CC exogenous compound on the expression of nucleic acid and amino acid
CC sequences of nucleic acid-associated proteins. The invention is useful in

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:02:57 ; Search time 45 Seconds
(without alignments)
141.467 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFIQSNLIRHQT.....KQCGKAFGCPNLRHRGRTH 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5 COMB pep.*
2: /cgn2_6/ptodata/1/1aa/6 COMB pep.*
3: /cgn2_6/ptodata/1/1aa/H COMB pep.*
4: /cgn2_6/ptodata/1/1aa/PCTUS COMB pep.*
5: /cgn2_6/ptodata/1/1aa/RE COMB pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	315	71.3	530	US-10-104-047-3434	Sequence 3434, Ap
2	315	71.3	656	US-09-949-016-10075	Sequence 10075, A
3	315	71.3	1191	US-09-949-016-6356	Sequence 6356, Ap
4	314	71.0	292	US-09-538-092-1334	Sequence 1334, Ap
5	309	69.9	321	US-09-538-092-1304	Sequence 1304, Ap
6	309	69.9	321	US-09-949-016-6364	Sequence 6364, Ap
7	309	69.9	334	US-09-949-016-8516	Sequence 8516, Ap
8	309	69.9	409	US-10-104-047-2663	Sequence 2663, Ap
9	309	69.9	475	US-10-104-047-3470	Sequence 3470, Ap
10	308	69.7	540	US-10-104-047-3748	Sequence 3748, Ap
11	306	69.2	286	US-10-104-047-2854	Sequence 2854, Ap
12	306	69.2	803	US-09-538-092-1026	Sequence 1026, Ap
13	305	69.0	586	US-10-104-047-2592	Sequence 2592, Ap
14	302	68.3	1805	US-09-494-190-125	Sequence 125, App
15	301	68.1	504	US-10-104-047-2053	Sequence 2053, Ap
16	301	68.1	575	US-10-104-047-3423	Sequence 3423, Ap
17	300	67.9	185	US-09-494-190-126	Sequence 126, App
18	299	67.6	325	US-09-538-092-1386	Sequence 1386, Ap
19	299	67.6	350	US-10-104-047-2500	Sequence 2500, Ap
20	299	67.6	469	US-09-538-092-1195	Sequence 1195, Ap
21	299	67.6	746	US-09-949-016-11494	Sequence 11494, A
22	298	67.4	685	US-09-949-016-8627	Sequence 8627, Ap
23	297	67.2	487	US-10-104-047-2868	Sequence 2868, Ap
24	293	66.3	367	US-09-949-016-6722	Sequence 6722, Ap
25	293	66.3	368	US-08-933-750C-14	Sequence 14, Appl
26	293	66.3	368	US-09-234-613-14	Sequence 14, Appl
27	293	66.3	368	US-09-949-016-6775	Sequence 6775, Ap

ALIGNMENTS

RESULT 1

US-10-104-047-3434
; Sequence 3434, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3434
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3434

Query Match 71.3%; Score 315; DB 2; Length 530;
Best Local Similarity 71.4%; Pred No. 98-29; Indels 0; Gaps 0;
Matches 55; Conservative 5; Mismatches 17;

QY 1 CKDCGKAFIQSNLIRHQTHTGKPYKCEGKAFQTSSNLTKEHKIHTGKPYKQCC 60
Db 412 CSECGKAFIQSNLIRHQTHTGKPYKCEGKAFQTSSNLTKEHKIHTGKPYHCNC 471

QY 61 GKAFGCPNLRHRGRTH 77
Db 472 GKAFGCPNLRHRGRTH 488

RESULT 2

US-09-949-016-10075
; Sequence 10075, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10075
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Human
.US-09-949-016-10075

Query Match          71.3%; Score 315; DB 2; Length 656;
Best Local Similarity 72.7%; Pred. No. 1.1e-28;
Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 243 CKDCGKAFIWSLTLRHRIHTGKPYKCECGKAFSRSSTLTQKHKIHTGKPYKCKEC 302

Qy 61 GKAFGCPNLRHGRTH 77
Db 303 GKAFKHSSALAKHKIIH 319

RESULT 3
US-09-949-016-6356
; Sequence 6356, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6356
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6356

Query Match          71.3%; Score 315; DB 2; Length 1191;
Best Local Similarity 72.7%; Pred. No. 2.2e-28;
Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 772 CKDCGKAFIWSLTLRHRIHTGKPYKCECGKAFSRSSTLTQKHKIHTGKPYKCKEC 831

Qy 61 GKAFGCPNLRHGRTH 77
Db 832 GKAFKHSSALAKHKIIH 848

RESULT 4
US-09-538-092-1334
; Sequence 1334, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/178,965

Query Match          69.9%; Score 309; DB 2; Length 321;
Best Local Similarity 66.2%; Pred. No. 2.6e-28;
Matches 51; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 126 CSDCGKAFIQSSILKGMRSHTGKPYKCECDHCGKFSQSSHLNVHVKRTHHTGKPYDKCEC 185

Qy 61 GKAFGCPNLRHGRTH 77
Db 186 GKFTVFSLLQKVRTH 202

RESULT 6
US-09-949-016-6364
; Sequence 6364, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

RESULT 8
US-10-104-047-2663
; Sequence 2663, Application US/10104047
; Patent No. 6943241

```

RESULT 10
US-10-104-047-3748
; Sequence 3748, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:

```

```
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3748
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3748

Query Match      69.7%; Score 308; DB 2; Length 540;
Best Local Similarity 67.5%; Pred. No. 6.1e-28;
Matches 52; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 286 CEECGKAFQTSSTLTGKRIHTGKPYKCEGCGKAFNRSSKLTGKHNHTGKPYKCEC 345

Qy 61 GKAFGCPNLRHGRTH 77
Db 346 GKAFNRSSNLTGHRKIH 362

RESULT 11
US-10-104-047-2854
; Sequence 2854, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2854
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2854

Query Match      69.2%; Score 306; DB 2; Length 286;
Best Local Similarity 64.9%; Pred. No. 5.2e-28;
Matches 50; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 40 CKQCGKAFSWSSVRIHERHTGKPYKCEGCGKSNFSSFRHRHTGKPYKCEC 99

Qy 61 GKAFGCPNLRHGRTH 77
Db 100 GKAFNCPSSFRHERTH 116

RESULT 12
US-09-538-092-1026
; Sequence 1026, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormatter Version 0.9
; SEQ ID NO 1026

; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3748
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3748

Query Match      69.7%; Score 308; DB 2; Length 540;
Best Local Similarity 67.5%; Pred. No. 6.1e-28;
Matches 52; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 286 CEECGKAFQTSSTLTGKRIHTGKPYKCEGCGKAFNRSSKLTGKHNHTGKPYKCEC 345

Qy 61 GKAFGCPNLRHGRTH 77
Db 346 GKAFNRSSNLTGHRKIH 362

RESULT 11
US-10-104-047-2854
; Sequence 2854, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2854
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2854

Query Match      69.2%; Score 306; DB 2; Length 286;
Best Local Similarity 64.9%; Pred. No. 5.2e-28;
Matches 50; Conservative 12; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 40 CKQCGKAFSWSSVRIHERHTGKPYKCEGCGKSNFSSFRHRHTGKPYKCEC 99

Qy 61 GKAFGCPNLRHGRTH 77
Db 100 GKAFNCPSSFRHERTH 116

RESULT 12
US-09-538-092-1026
; Sequence 1026, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormatter Version 0.9
; SEQ ID NO 1026

; LENGTH: 803
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P28160
US-09-538-092-1026

Query Match      69.2%; Score 306; DB 2; Length 803;
Best Local Similarity 71.4%; Pred. No. 1.6e-27;
Matches 55; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 534 CEECGKAFNHFSLITGKRIHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCEC 593

Qy 61 GKAFGCPNLRHGRTH 77
Db 594 GKAFQTQSSNLTGKHKIH 610

RESULT 13
US-10-104-047-2592
; Sequence 2592, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2592
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2592

Query Match      69.0%; Score 305; DB 2; Length 586;
Best Local Similarity 68.8%; Pred. No. 1.5e-27;
Matches 53; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 371 CDECGKAFNQSSLTGKRIHTGKPYKCEGCGKAFQSSLTGKHKIHTGKPYKCEC 430

Qy 61 GKAFGCPNLRHGRTH 77
Db 431 GKAFSSWSSAFTHKRNH 447

RESULT 14
US-09-494-190-125
; Sequence 125, Application US/09494190
; Patent No. 6610512
; GENERAL INFORMATION:
; APPLICANT: BARBAS, Carlos F.
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN
; FILE REFERENCE: TSRI 645.2
; CURRENT APPLICATION NUMBER: US/09/494,190
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: EP/99/07742
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/173,941
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 125
; LENGTH: 185
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:expressed
US-09-494-190-125

Query Match      68.3%; Score 302; DB 2; Length 185;
Best Local Similarity 63.6%; Pred. No. 9.4e-28;
Matches 49; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPKYKCECGKAFQTQSSNLTKKKKIHTGKPKYKCKOC 60
   |::|||::|:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 15 CPEGKSFGRKDSLVRHQRTHTGKPKYKCECGKAFQTQSSNLTKKKKIHTGKPKYKCKOC 60
   |::|||::|:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 61 GKAFGCPNLRHGRTH 77
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 75 GKSFSDCRDLARHQRTH 91
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```
RESULT 15
US-10-104-047-2053
; Sequence 2053, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2053
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2053
```

```
Query Match      68.1%; Score 301; DB 2; Length 504;
Best Local Similarity 67.5%; Pred. No. 3.7e-27;
Matches 52; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPKYKCECGKAFQTQSSNLTKKKKIHTGKPKYKCKOC 60
   |::|||::|:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 395 CEECGSAFRAFSTLTGKRVHTGKPKYKCECGKAFNWSSTLTKKRIHTGKPKYKCEC 454
   |::|||::|:|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 61 GKAFGCPNLRHGRTH 77
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 455 GKAFNRSSNLTTRKKIH 471
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Search completed: January 9, 2006, 20:10:34
Job time : 45 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:03:37 ; Search time 61 Seconds
(without alignments)
527.424 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFIQKSLNRHQT.....KQCGKAFGCPNLRHRGTH 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap:*
4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pap:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	378	4	US-10-314-669-2
2	442	100.0	378	4	US-10-669-861-2
3	374	84.6	209	4	US-10-314-669-14
4	374	84.6	209	4	US-10-669-861-14
5	340	76.9	107	5	US-10-746-864-60
6	340	76.9	111	5	US-10-732-620-57
7	334	75.6	111	5	US-10-732-620-67
8	330	74.7	111	5	US-10-732-620-56
9	328	74.2	111	5	US-10-732-620-58
10	328	74.2	111	5	US-10-732-620-58
11	328	74.2	113	5	US-10-732-620-44
12	327	74.0	555	5	US-10-450-763-53101
13	327	74.0	673	4	US-10-108-260A-4868
14	326	73.8	211	4	US-10-314-669-264
15	325	73.5	107	5	US-10-746-864-54
16	325	73.5	184	3	US-09-765-555-42
17	321	72.6	184	3	US-09-765-555-39
18	320	72.4	111	5	US-10-732-620-60
19	319	72.2	592	4	US-10-108-260A-3669
20	319	72.2	632	5	US-10-450-763-48745
21	317	71.7	111	5	US-10-732-620-46
22	316	71.5	99	4	US-09-911-261A-8
23	316	71.5	99	4	US-10-057-408-8
24	316	71.5	99	4	US-10-333-487-8
25	316	71.5	113	5	US-10-732-620-36
26	315	71.3	107	4	US-10-314-669-23
27	315	71.3	107	4	US-10-669-861-23

28	315	71.3	209	4	US-10-314-669-10	Sequence 10, Appl
29	315	71.3	209	4	US-10-669-861-10	Sequence 10, Appl
30	315	71.3	499	4	US-10-108-260A-2954	Sequence 2954, Ap
31	315	71.3	530	4	US-10-104-047-3434	Sequence 3434, Ap
32	315	71.3	1191	4	US-10-408-765A-2105	Sequence 2105, Ap
33	315	71.3	1212	5	US-10-450-763-30758	Sequence 30758, A
34	314	71.0	111	5	US-10-732-620-33	Sequence 33, Appl
35	314	71.0	111	5	US-10-732-620-39	Sequence 39, Appl
36	314	71.0	111	5	US-10-732-620-62	Sequence 62, Appl
37	312	70.6	183	5	US-10-732-620-27	Sequence 27, Appl
38	312	70.6	111	5	US-10-732-620-59	Sequence 59, Appl
39	312	70.6	184	3	US-09-765-555-38	Sequence 38, Appl
40	312	70.6	184	3	US-09-765-555-75	Sequence 75, Appl
41	311	70.4	99	3	US-09-911-261A-9	Sequence 9, Appl
42	311	70.4	99	3	US-09-911-261A-10	Sequence 10, Appl
43	311	70.4	99	4	US-10-057-408-9	Sequence 9, Appl
44	311	70.4	99	4	US-10-057-408-10	Sequence 10, Appl
45	311	70.4	99	4	US-10-333-487-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-10-314-669-2
; Sequence 2, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-314-669-2

Query Match	100.0%;	Score 442;	DB 4;	Length 378;
Best Local Similarity	100.0%;	Pred. No. 9.8e-39;		
Matches	77;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	CKDCGKAFIQKSLNRHQTHTGEPYKCEGKGAFTOSSNLTKHKIHTGKPKYKQKC	60	
Db	33	CKDCGKAFIQKSLNRHQTHTGEPYKCEGKGAFTOSSNLTKHKIHTGKPKYKQKC	92	
QY	61	GKAFGCPNLRHRGTH	77	
Db	93	GKAFGCPNLRHRGTH	109	
RESULT 2				

```
US-10-669-861-2
; Sequence 2, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Lee, Yangsoon
; APPLICANT: Kim, Jin-Soo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-669-861-2

Query Match 100.0%; Score 442; DB 4; Length 378;
Best Local Similarity 100.0%; Pred. No. 9.8e-39;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRHTGKPYKCEGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 33 CKDCGKAFIQKSNLIRHQRHTGKPYKCEGKAFQTQSSNLTGKHKIHTGKPYKCKQC 92

QY 61 GKAFGCPNLRHGRTH 77
Db 93 GKAFGCPNLRHGRTH 109

RESULT 3
US-10-314-669-14
; Sequence 14, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14

US-10-669-861-2
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-669-861-2

Query Match 84.6%; Score 374; DB 4; Length 209;
Best Local Similarity 79.2%; Pred. No. 9e-32;
Matches 61; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRHTGKPYKCEGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 61 CSECGRGFSQKSNLIHQRTHTGKPYKCEGKAFQTQSSNLTGKHKIHTGKPYKCKQC 120

QY 61 GKAFGCPNLRHGRTH 77
Db 121 GKAFGCPNLRHGRTH 137

RESULT 4
US-10-669-861-14
; Sequence 14, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Lee, Yangsoon
; APPLICANT: Kim, Jin-Soo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-669-861-14

Query Match 84.6%; Score 374; DB 4; Length 209;
Best Local Similarity 79.2%; Pred. No. 9e-32;
Matches 61; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRHTGKPYKCEGKAFQTQSSNLTGKHKIHTGKPYKCKQC 60
Db 61 CSECGRGFSQKSNLIHQRTHTGKPYKCEGKAFQTQSSNLTGKHKIHTGKPYKCKQC 120

QY 61 GKAFGCPNLRHGRTH 77
Db 121 GKAFGCPNLRHGRTH 137

RESULT 5
US-10-746-864-60
; Sequence 60, Application US/10746864
; Publication No. US20040259258A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Jang, Young-Soon
```

```

; TITLE OF INVENTION: REGULATION OF PROKARYOTIC GENE
; TITLE OF INVENTION: EXPRESSION WITH ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-012001
; CURRENT APPLICATION NUMBER: US/10/746,864
; CURRENT FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: US 10/669,861
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 164
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-746-864-60

```

```

Query Match 76.9%; Score 340; DB 5; Length 107;
Best Local Similarity 74.0%; Pred. No. 1.9e-28;
Matches 57; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQSSNLTQSKKIHHTGKPYKCKQC 60
Db 31 CDHCGKAFSVSSNLRVHRIHTGKPFCKDCGKAFIQKSNLIRHQRTHTGKPYKCKQC 90
QY 61 GKAFGCPNLRHRGRTH 77
Db 91 GKAFGCPNLRHRGRTH 107

```

```

RESULT 6
US-10-732-620-57
; Sequence 57, Application US/10732620
; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-57

```

```

Query Match 76.9%; Score 340; DB 5; Length 111;
Best Local Similarity 74.0%; Pred. No. 1.9e-28;
Matches 57; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQSSNLTQSKKIHHTGKPYKCKQC 60
Db 3 CKTCQKFSRSDHLKTHTRHTGKPYKCEGCGKAFNRSHLTRHRIHTGKPYKCKQC 62
QY 61 GKAFGCPNLRHRGRTH 77
Db 63 GKAFGCPNLRHRGRTH 79

```

```

RESULT 7
US-10-732-620-67
; Sequence 67, Application US/10732620
; Publication No. US20050032186A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-67

```

```

Query Match 75.6%; Score 334; DB 5; Length 111;
Best Local Similarity 75.3%; Pred. No. 8.4e-28;
Matches 58; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQSSNLTQSKKIHHTGKPYKCKQC 60
Db 31 CPDCGKFSQSSSLRHRHQRTHTGKPYKCEGCGKAFQSSNLTQSKKIHHTGKPYKCKQC 90
QY 61 GKAFGCPNLRHRGRTH 77
Db 91 GKAFNRSHLTRHQRH 107

```

```

RESULT 8
US-10-732-620-56
; Sequence 56, Application US/10732620
; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732,620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-732-620-56

```

```

Query Match 74.7%; Score 330; DB 5; Length 111;
Best Local Similarity 71.4%; Pred. No. 2.2e-27;
Matches 55; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGCGKAFQSSNLTQSKKIHHTGKPYKCKQC 60
Db 31 CRKCGRFSRSLRHRHQRTHTGKPYKCEGCGKAFNRSHLTRHRIHTGKPYKCKQC 90
QY 61 GKAFGCPNLRHRGRTH 77
Db 91 GKAFRWPNSLTRHKEIH 107

```

```

RESULT 9
US-10-732-620-38
; Sequence 38, Application US/10732620

```

; Publication No. US20050032186A1

; GENERAL INFORMATION:

; APPLICANT: Kim, Jin-Soo

; APPLICANT: Shin, Hyun-Chul

; APPLICANT: Kwon, Heung-Sun

; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS

; FILE REFERENCE: 12279-009001

; CURRENT APPLICATION NUMBER: US/10/732,620

; CURRENT FILING DATE: 2003-12-09

; PRIOR APPLICATION NUMBER: US 60/431,892

; PRIOR FILING DATE: 2002-12-09

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 38

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetically generated peptide

US-10-732-620-38

Query Match 74.2%; Score 328; DB 5; Length 111;

Best Local Similarity 71.4%; Pred. No. 3.6e-27;

Matches 55; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGKAFCTOSSNLTKHKIHTGKPYKCKQC 60

Db 31 CSDCGKAFRDSCLNRRHRTHTGKPYKCEGKAFCTOSSNLTKHKIHTGKPYKCKQC 90

Qy 61 GKAFGCPNLRHGRTH 77

Db 91 GKAFGCPNLRHGRTH 107

RESULT 10

US-10-732-620-58

; Sequence 58, Application US/10732620

; Publication No. US20050032186A1

; GENERAL INFORMATION:

; APPLICANT: Kim, Jin-Soo

; APPLICANT: Shin, Hyun-Chul

; APPLICANT: Kwon, Heung-Sun

; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS

; FILE REFERENCE: 12279-009001

; CURRENT APPLICATION NUMBER: US/10/732,620

; CURRENT FILING DATE: 2003-12-09

; PRIOR APPLICATION NUMBER: US 60/431,892

; PRIOR FILING DATE: 2002-12-09

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 58

; LENGTH: 111

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetically generated peptide

US-10-732-620-58

Query Match 74.2%; Score 328; DB 5; Length 111;

Best Local Similarity 55.2%; Pred. No. 3.6e-27;

Matches 58; Conservative 10; Mismatches 9; Indels 28; Gaps 1;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGKAFCTOSSNLTKHKIHTGKPYKCKQC 62

Db 3 CMECKAFNRSHLRHQRTHTGKPYKCEGKAFCTOSSNLTKHKIHTGKPYKCKQC 35

Qy 36 ---FTQSSNLTKHKIHTGKPYKCKQCQKAFGCPNLRHGRTH 77

Db 63 QKFSRSDHLKTHRTHTGKPYKCKQCQKAFGCPNLRHGRTH 107

RESULT 11

US-10-732-620-44

; Sequence 44, Application US/10732620

; Publication No. US20050032186A1

; GENERAL INFORMATION:

; APPLICANT: Kim, Jin-Soo

; APPLICANT: Shin, Hyun-Chul

; APPLICANT: Kwon, Heung-Sun

; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS

; FILE REFERENCE: 12279-009001

; CURRENT APPLICATION NUMBER: US/10/732,620

; CURRENT FILING DATE: 2003-12-09

; PRIOR APPLICATION NUMBER: US 60/431,892

; PRIOR FILING DATE: 2002-12-09

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 44

; LENGTH: 113

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetically generated peptide

US-10-732-620-44

Query Match 74.2%; Score 328; DB 5; Length 113;

Best Local Similarity 75.3%; Pred. No. 3.7e-27;

Matches 58; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 1 CKDCGKAFIQKSNLIRHQRTHTGKPYKCEGKAFCTOSSNLTKHKIHTGKPYKCKQC 60

Db 33 CPDCGKFSOSSLIRHQRTHTGKPYKCEGKAFCTOSSNLTKHKIHTGKPYKCKQC 92

Qy 61 GKAFGCPNLRHGRTH 77

Db 93 GRGFSRKSNIHQRTH 109

RESULT 12

US-10-450-763-53101

; Sequence 53101, Application US/10450763

; Publication No. US20050196754A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 790CIP3/US

; CURRENT APPLICATION NUMBER: US/10/450,763

; CURRENT FILING DATE: 2003-06-11

; PRIOR APPLICATION NUMBER: PCT/US01/08631

; PRIOR FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: 09/540,217

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/649,167

; PRIOR FILING DATE: 2000-08-23

; NUMBER OF SEQ ID NOS: 60736

; SOFTWARE: Custom

; SEQ ID NO 53101

; LENGTH: 555

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: DOMAIN

; LOCATION: (457)..(510)

; OTHER INFORMATION: S-adenosyl-L-homocysteine hydrolase proteins domain

; OTHER INFORMATION: identified by eMATRIX, accession number BL00738H, p-value=2.301e-

; OTHER INFORMATION: 17, raw score of 23.08

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (85)..(437)

; OTHER INFORMATION: zinc finger, C2H2 type domain identified by Pfam, accession

; OTHER INFORMATION: name zf-C2H2, E-value=3.9e-42, Pfam score of 153.4

US-10-450-763-53101

Query Match 74.0%; Score 327; DB 5; Length 555;

Best Local Similarity 72.7%; Pred. No. 2.4e-26;

Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:06:02 ; Search time 8 Seconds
(without alignments)
81.751 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCGKAFIQKSNLIRHQT.....KQCGKAFGCPNLRHRGRTH 77

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 61141 seqs, 8493638 residues

Total number of hits satisfying chosen parameters: 61141

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11 NEW PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	289	65.4	432	7 US-11-000-463-365	Sequence 365, App
2	281	63.6	100	7 US-11-115-922-223	Sequence 223, App
3	208	47.1	135	7 US-11-108-172-202	Sequence 202, App
4	202	45.7	310	7 US-11-148-149-1	Sequence 1, Appli
5	197	44.6	449	7 US-11-071-062-1	Sequence 1, Appli
6	197	44.6	449	7 US-11-196-459-1	Sequence 1, Appli
7	197	44.6	449	7 US-11-196-459-2	Sequence 2, Appli
8	194	43.9	429	7 US-11-071-062-9	Sequence 9, Appli
9	193	43.7	94	7 US-11-115-922-221	Sequence 221, App
10	176	39.8	578	6 US-10-821-234-1039	Sequence 1039, Ap
11	143	32.4	54	7 US-11-196-670-62	Sequence 62, Appl
12	140	31.7	54	7 US-11-196-670-64	Sequence 64, Appl
13	133	30.1	644	7 US-11-063-343-21	Sequence 21, Appl
14	131	29.6	52	7 US-11-196-670-61	Sequence 61, Appl
15	131	29.6	52	7 US-11-196-670-63	Sequence 63, Appl
16	117	26.5	466	7 US-11-110-082-33	Sequence 33, Appl
17	111.5	25.2	530	7 US-11-110-082-32	Sequence 32, Appl
18	93	21.0	496	7 US-11-110-082-31	Sequence 31, Appl
19	88	19.9	227	7 US-11-000-463-301	Sequence 301, App
20	86	19.5	28	7 US-11-196-670-54	Sequence 54, Appl
21	85	19.2	30	7 US-11-096-706-212	Sequence 212, App
22	85	19.2	30	7 US-11-101-287-157	Sequence 157, App
23	85	19.2	30	7 US-11-115-922-217	Sequence 217, App
24	79	17.9	28	7 US-11-096-706-213	Sequence 213, App
25	79	17.9	28	7 US-11-101-287-158	Sequence 158, App

Sequence 218, App
Sequence 58, Appl
Sequence 56, Appl
Sequence 57, Appl
Sequence 59, Appl
Sequence 5858, Ap
Sequence 60, Appl
Sequence 10, Appl
Sequence 9, Appli
Sequence 214, App
Sequence 159, App
Sequence 219, App
Sequence 296, App
Sequence 1033, Ap
Sequence 1034, Ap
Sequence 8036, Ap
Sequence 2, Appli
Sequence 55, Appl
Sequence 1011, Ap
Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-11-000-463-365
; Sequence 365, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 365
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-365

Query Match 65.4%; Score 289; DB 7; Length 432;
Best Local Similarity 62.3%; Pred. No. 8.5e-26;
Matches 48; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQTGTGKPCGKAFQTQSSNLTKHKIHTGKPKYKQC 60
||:|||||: : ||||| ||||| : ||::||| ||||| : |||

Query Match 43.9%; Score 194; DB 7; Length 429;
Best Local Similarity 43.4%; Pred. No. 4.7e-15;
Matches 33; Conservative 16; Mismatches 25; Indels 2; Gaps 1;

Qy 4 CGKAFIQKSLNRHQRTHGKPYKCE--ECGKAFQSSNLTGKHKIHTGKPYKCKQCG 61
Db 313 CNKRYFKLSHLQMSRKHTGKPYQCDKCFRRFRSDQLKRHRHTGKPYKCKQCG 372

Qy 62 KAFGCPNLRHGRTH 77
Db 373 RKFRSDHLKTHTRTH 388

RESULT 9
US-11-115-922-221
; Sequence 221, Application US/11115922
; Publication No. US20050267062A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo BioSciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT FILING DATE: 2005-04-26
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 221
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sp-1 transcription factor
US-11-115-922-221

Query Match 43.7%; Score 193; DB 7; Length 94;
Best Local Similarity 46.1%; Pred. No. 1.3e-15;
Matches 35; Conservative 13; Mismatches 26; Indels 2; Gaps 1;

Qy 4 CGKAFIQKSLNRHQRTHGKPYKCE--ECGKAFQSSNLTGKHKIHTGKPYKCKQCG 61
Db 14 CGKVGKTSHLRAHLRWHTGKPYKCE--ECGKAFQSSNLTGKHKIHTGKPYKCKQCG 73

Qy 62 KAFGCPNLRHGRTH 77
Db 74 KFRMSDHLKTHTRTH 89

RESULT 10
US-10-821-234-1039
; Sequence 1039, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1039
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1039

Query Match 39.8%; Score 176; DB 6; Length 578;
Best Local Similarity 43.4%; Pred. No. 7e-13;
Matches 33; Conservative 13; Mismatches 26; Indels 4; Gaps 2;

Qy 4 CGKAFIQKSLNRHQRTHGKPYKCECGKAFQSSNLTGKHKIHTGKPYKCKQCGKA 63
Db 357 CGRLFRLOKQLLRHAKHHTDORDYCEVCARAFKSHHNLAVHRMHTGKPLQCEICG-- 414

Qy 64 FGC--PSNLRHGRTH 77
Db 415 FTRQKASLNLNWMKKH 430

RESULT 11
US-11-196-670-62
; Sequence 62, Application US/11196670
; Publication No. US20050272090A1
; GENERAL INFORMATION:
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
; TITLE OF INVENTION: Nucleic Acid Binding of Multi Zinc Finger Transcription Factors
; FILE REFERENCE: 2676 5174US
; CURRENT APPLICATION NUMBER: US/11/196,670
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/028,396
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 99202068.5
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/EP00/05582
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: SIPiczf2+czf3
US-11-196-670-62

Query Match 32.4%; Score 143; DB 7; Length 54;
Best Local Similarity 51.1%; Pred. No. 3.2e-10;
Matches 23; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CKDCKGKAFIQKSLNRHQRTHGKPYKCECGKAFQSSNLTGKHKIHTGKPYKCKQCG 45
Db 1 CQICKKAFKHKHLLIEHSLRSLHSGKPYQCDKCKGRFSGHSYSQH 45

RESULT 12
US-11-196-670-64
; Sequence 64, Application US/11196670
; Publication No. US20050272090A1
; GENERAL INFORMATION:
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
; TITLE OF INVENTION: Nucleic Acid Binding of Multi Zinc Finger Transcription Factors
; FILE REFERENCE: 2676 5174US
; CURRENT APPLICATION NUMBER: US/11/196,670
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/028,396
; PRIOR FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: 99202068.5
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/EP00/05582
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: sigma-BFLczf2+czf3
US-11-196-670-64

Query Match 31.7%; Score 140; DB 7; Length 54;
Best Local Similarity 51.1%; Pred. No. 7e-10;
Matches 23; Conservative 9; Mismatches 13; Indels 0; Gaps 0;
QY 1 CKDCGKAFQKSNLIRHQRTHGKPYKCECGKAFQKSSNLTQH 45
Db 1 CGICKAFKHKHLEHMRHLSGKPYQCDKCGKRFSGSGYSQH 45

RESULT 13
US-11-063-343-21
; Sequence 21, Application US/11063343
; Publication No. US20050272090A1
; GENERAL INFORMATION:
; APPLICANT: Petroziello, Joseph M.
; APPLICANT: Carter, Paul
; TITLE OF INVENTION: Expression Profiling in Non-Small Cell
; FILE OF INVENTION: Lung Cancer
; FILE REFERENCE: 2681-1-003N
; CURRENT APPLICATION NUMBER: US/11/063,343
; CURRENT FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: 60/546,019
; PRIOR FILING DATE: 2004-02-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-063-343-21

Query Match 30.1%; Score 133; DB 7; Length 644;
Best Local Similarity 29.5%; Pred. No. 5.7e-08;
Matches 28; Conservative 14; Mismatches 33; Indels 20; Gaps 2;
QY 1 CKDCGKAFQKSNLIRH-----QRTHTGKPYKCECGKAFQKSSNL 42
Db 536 CPNCGQRFETENLVVHMSCLDQDMFKSAIMEENERDH--RRKHFCLNCGKGFYQRCFL 593
QY 43 TKHKIHTGKPYKQCKGAFGCPNLRHRGRTH 77
Db 594 REHYTVHTKEQPVQCTCGKQFLRLRLHNDMH 628

RESULT 14
US-11-196-670-61
; Sequence 61, Application US/11196670
; Publication No. US20050272090A1
; GENERAL INFORMATION:
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
; TITLE OF INVENTION: Nucleic Acid Binding of Multi Zinc Finger Transcription Factors
; FILE REFERENCE: 2676 5174US
; CURRENT APPLICATION NUMBER: US/11/196,670
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/028,396
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 99202068.5
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/EP00/05582

; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: SIPlnzf3+nzf4
US-11-196-670-61

Query Match 29.6%; Score 131; DB 7; Length 52;
Best Local Similarity 48.9%; Pred. No. 7e-09;
Matches 22; Conservative 7; Mismatches 16; Indels 0; Gaps 0;
QY 1 CKDCGKAFQKSNLIRHQRTHGKPYKCECGKAFQKSSNLTQH 45
Db 1 CTECGKAFKYKHLKHLRIHSGEKPYECPNCKKRFSGSGYSYSH 45

RESULT 15
US-11-196-670-63
; Sequence 63, Application US/11196670
; Publication No. US20050272090A1
; GENERAL INFORMATION:
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
; TITLE OF INVENTION: Nucleic Acid Binding of Multi Zinc Finger Transcription Factors
; FILE REFERENCE: 2676 5174US
; CURRENT APPLICATION NUMBER: US/11/196,670
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/028,396
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 99202068.5
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: PCT/EP00/05582
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: sigma-EFlnzf3+nzf4
US-11-196-670-63

Query Match 29.6%; Score 131; DB 7; Length 52;
Best Local Similarity 48.9%; Pred. No. 7e-09;
Matches 22; Conservative 7; Mismatches 16; Indels 0; Gaps 0;
QY 1 CKDCGKAFQKSNLIRHQRTHGKPYKCECGKAFQKSSNLTQH 45
Db 1 CTECGKAFKYKHLKHLRIHSGEKPYECPNCKKRFSGSGYSYSH 45

Search completed: January 9, 2006, 20:12:06
Job time : 8 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:00:51 ; Search time 39 Seconds
(without alignments)
189.966 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFIQKSLIRHQT.....KQCGKAFGCPNLRHGRTH 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	315	71.3	1191	2 S35305	zinc finger protei
2	314	71.0	292	2 S43826	finger protein OZF
3	311	70.4	169	2 A39240	finger protein mfg
4	310	70.1	427	2 A35659	krueppel-related p
5	307	69.5	595	2 G02075	transcription repr
6	306	69.2	229	2 A48927	Kruppel-like zinc
7	306	69.2	732	2 S47073	finger protein HZF
8	306	69.2	803	2 S26823	zinc finger protei
9	301	68.1	105	2 B48827	zinc finger protei
10	301	68.1	543	2 B34612	zinc finger protei
11	299	67.6	247	2 S47070	zinc finger protei
12	299	67.6	325	2 I38616	zinc finger protei
13	299	67.6	469	2 I38600	zinc finger protei
14	299	67.6	651	2 B32891	finger protein 2
15	299	67.6	686	2 A34612	zinc finger protei
16	295	66.7	519	2 A38073	transcription acti
17	294	66.5	488	2 S47072	finger protein HZF
18	294	66.5	555	2 A56560	zinc finger protei
19	294	66.5	564	2 T12489	hypothetical prote
20	293	66.3	319	2 T46469	hypothetical prote
21	293	66.3	347	2 S00549	developmental cont
22	292	66.1	93	2 PQ0636	zinc finger protei
23	292	66.1	614	2 JH0500	zinc finger protei
24	290	65.6	109	2 H45193	zinc finger protei
25	290	65.6	474	2 I54338	zinc finger protei
26	289	65.4	200	2 S47067	finger protein HZF
27	289	65.4	542	2 A54661	zinc finger protei
28	288	65.2	348	2 I38599	zinc finger protei
29	287	64.9	386	2 T12527	hypothetical prote

30	287	64.9	594	2 T12488	hypothetical prote
31	287	64.9	728	2 A48830	probable transcrip
32	286	64.7	194	2 I53859	zinc finger protei
33	285	64.5	462	2 S10397	finger protein kox
34	285	64.5	540	2 T14748	hypothetical prote
35	283	64.0	337	2 S60520	finger protein ZNF
36	283	64.0	411	2 S10245	finger protein, te
37	283	64.0	636	2 I48689	gene NK10 protein
38	283	64.0	693	2 I37570	zinc finger protei
39	282	63.8	378	2 S33994	finger protein ZNF
40	282	63.8	553	2 S22954	finger protein zfp
41	282	63.8	555	2 I53869	zinc finger protei
42	281	63.6	195	2 S00754	zinc finger protei
43	281	63.6	242	2 S30238	finger protein ZNF
44	281	63.6	261	2 S70006	finger protein zfo
45	281	63.6	393	2 JN0533	finger protein pml

ALIGNMENTS

RESULT 1

S35305
zinc finger protein ZNF91 - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: S35305
R:Beliefroid, E.J.; Marine, J.C.; Ried, T.; Lecocq, P.J.; Riviere, M.; Anemiy, C.; Pon
EMBO J. 12, 1363-1374, 1993
A>Title: Clustered organization of homologous KRAB zinc-finger genes with enhanced exp
A:Reference number: S35305; MUID:93223677; PMID:8467795
A:Accession: S35305
A:Molecule type: mRNA
A:Residues: 1-1191 <BEL>
A:Cross-references: UNIPROT:Q05481; UNIPARC:UPI000013C42A; EMBL:L11672; NID:g186773; P
A>Note: The authors translated the codon GCA for residue 750 as Thr and GCT for residue
C:Genetics:
A:Gene: GDB:ZNF91; HPF7; HTF10
A:Cross-references: GDB:132284
A:Map position: 19p12-19p12
C:Keywords: DNA binding; zinc finger

Query Match 71.3%; Score 315; DB 2; Length 1191;

Best Local Similarity 72.7%; Pred. No. 2e-23;

Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 C C C K A F I Q K S L I R H Q T K Q C G K A F G C P N L R H G R T H 77
DB 772 C K E C G K A F I Q K S L I R H Q T K Q C G K A F G C P N L R H G R T H 77
QY 61 G K A F G C P N L R H G R T H 77
DB 832 G K A F G C P N L R H G R T H 848

RESULT 2

S43826
finger protein OZF, Krueppel-related - human
C:Species: Homo sapiens (man)
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43826
J. Mol. Biol. 236, 399-404, 1994
A>Title: The OZF gene encodes a protein consisting essentially of zinc finger motifs.
A:Reference number: S43826; MUID:94149744; PMID:8107129
A:Accession: S43826
A:Molecule type: mRNA
A:Residues: 1-292 <LEG>
A:Cross-references: UNIPROT:Q15072; UNIPARC:UPI0000130F91; EMBL:X70394; NID:g468707; P
A:Gene: GDB:ZNF146; OZF
A:Cross-references: GDB:230289; OMIM:601505
A:Map position: 19q13.1-19q13.1

C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C;Keywords: DNA binding; zinc finger

Query Match 71.0%; Score 314; DB 2; Length 292;
Best Local Similarity 70.1%; Pred. No. 7.8e-24;
Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAFIOKSNLIRHQRTHGKPYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 60
102 CQCEGKAFNQSSNLTTHKKIHTGKPKYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 161
DB CQCEGKAFNQSSNLTTHKKIHTGKPKYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 161

QY 61 GKAFGCPNLRHGRTH 77
DB 162 GATGQKYLKHQNIH 178

RESULT 3
A39240
finger protein mfgl - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 09-Jul-2004
C;Accession: A39240
R;Passananti, C.; Felsani, A.; Caruso, M.; Amati, P.
Proc. Natl. Acad. Sci. U.S.A. 86, 9417-9421, 1989
A;Title: Mouse genes coding for "zinc-finger"-containing proteins: characterization and
A;Reference number: A39240; MUID:90083278; PMID:2512579
A;Accession: A39240
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-169 <PAS>
A;Cross-references: UNIPROT:P16372; UNIPARC:UPI0000029AEE; GB:M28513; NID:g199136; PIDN:
C;Keywords: DNA binding; zinc finger

Query Match 70.4%; Score 311; DB 2; Length 169;
Best Local Similarity 68.8%; Pred. No. 9.7e-24;
Matches 53; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 1 CKDCGKAFIOKSNLIRHQRTHGKPYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 60
DB 38 CEECGKAFNQSSNLTTHKKIHTGKPKYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 97

QY 61 GKAFGCPNLRHGRTH 77
DB 98 GKAFSTSSNLSHKKIH 114

RESULT 4
A35659
krueppel-related protein H-plk - human
C;Species: Homo sapiens (man)
C;Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 09-Jul-2004
C;Accession: A35659
R;Kato, N.; Shimotohno, K.; VanLeeuwen, D.; Cohen, M.
Mol. Cell. Biol. 10, 4401-4405, 1990
A;Title: Human proviral mRNAs down regulated in choriocarcinoma encode a zinc finger pro
A;Reference number: A35659; MUID:90318410; PMID:2115127
A;Accession: A35659
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <KAT>
A;Cross-references: UNIPROT:Q02313; UNIPARC:UPI000006E47C; GB:M55422; NID:g184342; PIDN:
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 70.1%; Score 310; DB 2; Length 427;
Best Local Similarity 71.4%; Pred. No. 2.7e-23;
Matches 55; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 1 CKDCGKAFIOKSNLIRHQRTHGKPYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 60
DB 279 CEECGKAFNQSSNLTTHKKIHTGKPKYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 338

QY 61 GKAFGCPNLRHGRTH 77

Db 339 GKAFNQLSNLRRHKVIH 355

RESULT 5
G02075
transcription repressor zinc finger protein 85 - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C;Accession: G02075
R;Poncellet, D.A.
submitted to the EMBL Data Library, September 1995
A;Reference number: G09169
A;Accession: G02075
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-595 <PON>
A;Cross-references: UNIPROT:Q03923; UNIPARC:UPI000013C428; EMBL:U35376; NID:g1017721; PI:
C;Genetics:
A;Gene: GDB:ZNF85
A;Cross-references: GDB:L32279
A;Map position: 19p12-19p12
C;Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 69.5%; Score 307; DB 2; Length 595;
Best Local Similarity 68.8%; Pred. No. 7e-23;
Matches 53; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 1 CKDCGKAFIOKSNLIRHQRTHGKPYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 60
DB 232 CEECGKAFNQSSNLTTHKKIHTGKPKYKCEGKGAFTQSSNLTTHKKIHTGKPKYCKQC 291

QY 61 GKAFGCPNLRHGRTH 77
DB 292 GKAFNRSSLTTHRRKIH 308

RESULT 6
A48927
Kruppel-like zinc finger protein HKR-T1 - human
N;Alternate names: zinc finger protein kok15; zinc finger protein ZNF22
C;Species: Homo sapiens (man)
C;Date: 19-Dec-1993 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: A48927; I37954; C56409; S10403
R;Wu, B.Y.; Hanley, E.W.; Turka, L.A.; Nabel, G.J.
Blood 80, 2571-2576, 1992
A;Title: Isolation of a cDNA clone encoding a zinc finger protein highly expressed in T-
A;Reference number: A48927; MUID:93043304; PMID:1421376
A;Accession: A48927
A;Molecule type: mRNA
A;Residues: 1-229 <WU1>
A;Cross-references: UNIPROT:P17026; UNIPARC:UPI00000493DA; GB:S50223; NID:g260311; PIDN:
A;Experimental source: MOLT 4 T-cells
A;Note: the authors translated the codon ATC for residue 76 as Asn, and the codon ACT for
A;Note: sequence modified after extraction from NCBI backbone (NCBIN:118663)
R;Thiesen, H.J.
New Biol. 2, 363-374, 1990

A;Title: Multiple genes encoding zinc finger domains are expressed in human T cells.
A;Reference number: I37949; MUID:91145339; PMID:2288909
A;Accession: I37954
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 139-194 <TH1>
A;Cross-references: UNIPARC:UPI000016ABBD; EMBL:X52346; NID:g34147; PIDN:CRA36572.1; PID:
A;Note: submitted to the EMBL Data Library, March 1990
R;Bray, P.; Lichter, P.; Thiesen, H.J.; Ward, D.C.; Dawid, I.B.
Proc. Natl. Acad. Sci. U.S.A. 88, 9563-9567, 1991
A;Title: Characterization and mapping of human genes encoding zinc finger proteins.
A;Reference number: A56409; MUID:92052132; PMID:1946370
A;Accession: C56409
A;Molecule type: DNA
A;Residues: 132-209, 'Y', 211-212, 'LVQE' <BRA>
A;Cross-references: UNIPARC:UPI0000178A38; GB:M771172; NID:g340429
A;Note: the nucleotide sequence and translation in GenBank entry HUMZFPAC, release 117.0

A/Cross-References: UNIPROT:PI17038; UNIPARC:UPI000013C415; EMBL:X59244; NID:g38031; PIDN

A;Status: preliminary

New Biol. 2, 363-374, 1990
A:Title: Multiple genes encoding zinc finger domains are expressed in human T cells.
A:Reference number: I37949; MUID:91145339; PMID:2288909
A:Accession: I37572
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2006, 20:00:06 ; Search time 160 Seconds
(without alignments)
339.535 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109
Perfect score: 442
Sequence: 1 CKDCGKAFQKSNLRHQRT.....KQCGKAFGCPNLRHRGRT 77

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	74.0	638	2 Q86VX3 HUMAN	Q86vx3 homo sapien
2	327	74.0	673	1 ZN433 HUMAN	Q8n7k0 homo sapien
3	319	72.2	592	2 Q8N8Q4 HUMAN	Q8n8q4 homo sapien
4	319	72.2	616	1 ZNF93 HUMAN	P35789 homo sapien
5	315	71.3	173	2 Q9H9X5 HUMAN	Q9h9x5 homo sapien
6	315	71.3	403	2 Q59FB0 HUMAN	Q59fb0 homo sapien
7	315	71.3	503	2 Q4G180 HUMAN	Q4g180 homo sapien
8	315	71.3	530	2 Q8NA64 HUMAN	Q8na64 homo sapien
9	315	71.3	563	2 Q6ZMR0 HUMAN	Q6zmr0 homo sapien
10	315	71.3	1191	1 ZNF91 HUMAN	Q05481 homo sapien
11	314	71.0	292	1 OZF HUMAN	Q15072 homo sapien
12	314	71.0	292	2 Q5RAA6 PONPY	Q5raa6 pongo pygma
13	314	71.0	292	2 Q5RFP4 PONPY	Q5rfp4 pongo pygma
14	314	71.0	544	1 ZNS02 HUMAN	Q8tbz5 homo sapien
15	314	71.0	589	2 Q6AYV1 RAT	Q6ayv1 rattus norv
16	314	71.0	611	2 Q8CII6 MOUSE	Q8cii6 mus musculu
17	313	70.8	292	1 OZF BOVIN	Q28151 bos taurus
18	313	70.8	292	2 Q9Z0Q5 MOUSE	Q9z0q5 mus musculu
19	313	70.8	302	2 Q8BQK6 MOUSE	Q8bqn6 mus musculu
20	313	70.8	584	2 Q86XU0 HUMAN	Q86xu0 homo sapien
21	313	70.8	626	2 Q99KB9 MOUSE	Q99kb9 mus musculu
22	313	70.8	653	2 Q5DTH6 MOUSE	Q5dth6 mus musculu
23	312	70.6	395	2 Q43693 HUMAN	Q43693 homo sapien
24	311	70.4	169	1 ZFP58 MOUSE	P16372 mus musculu
25	311	70.4	338	2 Q8NC79 HUMAN	Q8nc79 homo sapien
26	311	70.4	530	2 Q8NEM1 HUMAN	Q8nem1 homo sapien
27	311	70.4	536	2 Q96WL5 HUMAN	Q96wl5 homo sapien
28	311	70.4	588	1 ZN429 HUMAN	Q86v71 homo sapien
29	311	70.4	688	1 ZN429 HUMAN	Q8taq5 homo sapien
30	311	70.4	783	1 ZNS88 HUMAN	Q9uii5 homo sapien
31	310	70.1	427	2 Q02313 HUMAN	Q02313 homo sapien

32	310	70.1	481	2 Q4R3H8 MACFA	Q4r3h8 macaca fasc
33	310	70.1	529	2 Q5FWH5 MOUSE	Q5fwh5 mus musculu
34	310	70.1	546	2 Q7M6X4 MOUSE	Q7m6x4 mus musculu
35	309	69.9	238	2 Q8BYD9 MOUSE	Q8byd9 mus musculu
36	309	69.9	321	1 ZN177 HUMAN	Q13360 homo sapien
37	309	69.9	402	2 Q8NCA3 HUMAN	Q8nc43 homo sapien
38	309	69.9	423	2 Q6POL2 HUMAN	Q6pol2 homo sapien
39	309	69.9	471	2 Q8NCF9 HUMAN	Q8ncf9 homo sapien
40	309	69.9	475	2 Q8NA42 HUMAN	Q8na42 homo sapien
41	309	69.9	475	2 Q6X2C7 HUMAN	Q6x2c7 homo sapien
42	309	69.9	475	2 Q4R6C3 MACFA	Q4r6c2 macaca fasc
43	309	69.9	519	2 Q96M04 HUMAN	Q96m04 homo sapien
44	308	69.7	502	2 Q59F60 HUMAN	Q59f60 homo sapien
45	308	69.7	540	2 Q8N211 HUMAN	Q8n211 homo sapien

ALIGNMENTS

RESULT 1
Q86VX3 HUMAN PRELIMINARY; PRT; 638 AA.
ID Q86VX3; HUMAN
AC Q86VX3;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE ZNF433 protein.
GN Name=ZNF433;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsu L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Griinwood J., Schmutz J., Myers R.M.,
RA Butterfield A.C., Griinwood J., Schmutz J., Myers R.M.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Testis;
RG NIH MGC Project;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC047412; AAH47412.1; -; mRNA.
DR HSSP; P08047; 1SP2.
DR Ensembl; ENSG00000197647; Homo sapiens.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 19.
DR PRINTS; PR00048; ZINCfinger.

```

DR ProDom; PD000003; Znf C2H2; 16.
DR SMART; SM00355; Znf C2H2; 19.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 18.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 19.
DR Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 638 AA; 73196 MW; 4E05BFC8FB53F1FD CRC64;

Query Match          74.0%; Score 327; DB 2; Length 638;
Best Local Similarity 72.7%; Pred. No. 2.2e-26;
Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGKPYKCEGKAFQSSNLTQSKHKKIHTGKPKYCKQC 60
Db 529 CKQCGKAFGSAHLMQHGRTHTGKPYCKQCGKSGFCASRLQMHGRTHGKPKYCKQC 588
QY 61 GKAFGCPNLRHGRTH 77
Db 589 GKAFGCPNLRHGRTH 605

RESULT 2
ZN433 HUMAN
ID ZN433_HUMAN STANDARD; PRT; 673 AA.
AC Q8N7K0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zinc finger protein 433.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Uterus;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Wajima T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo M.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Mizoguchi S., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsuura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RL Nat. Genet. 36:40-45(2004).
CC -!- FUNCTION: May be involved in transcriptional regulation.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
CC family.

```

```

CC -!- SIMILARITY: Contains 19 C2H2-type zinc fingers.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AK098300; BAC05279.1; -; mRNA.
CC HSSP; P08047; 1SP2.
CC DR Ensembl; ENSG00000197647; Homo sapiens.
CC DR HGNC; HGNC:20811; ZNF433.
CC DR InterPro; IPR001909; KRAB.
CC DR InterPro; IPR007087; Znf_C2H2.
CC DR InterPro; IPR007086; Znf_C2H2_sub.
CC DR Pfam; PF01352; KRAB; 1.
CC DR Pfam; PF00096; zf-C2H2; 19.
CC DR PRINTS; PR00048; ZINCFINGER.
CC DR ProDom; PD000003; Znf C2H2; 16.
CC DR SMART; SM00349; KRAB; 1.
CC DR SMART; SM00355; Znf C2H2; 19.
CC DR PROSITE; PS50805; KRAB; 1.
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 18.
CC DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 19.
CC KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Transcription;
CC Transcription regulation; Zinc; Zinc-finger.
CC DOMAIN 7 85
CC FT ZN_FING 142 164 C2H2-type 1.
CC FT ZN_FING 170 192 C2H2-type 2.
CC FT ZN_FING 198 220 C2H2-type 3.
CC FT ZN_FING 226 248 C2H2-type 4.
CC FT ZN_FING 254 276 C2H2-type 5.
CC FT ZN_FING 282 304 C2H2-type 6.
CC FT ZN_FING 310 332 C2H2-type 7.
CC FT ZN_FING 338 360 C2H2-type 8.
CC FT ZN_FING 366 388 C2H2-type 9.
CC FT ZN_FING 394 416 C2H2-type 10.
CC FT ZN_FING 422 444 C2H2-type 11.
CC FT ZN_FING 450 472 C2H2-type 12.
CC FT ZN_FING 478 500 C2H2-type 13.
CC FT ZN_FING 506 528 C2H2-type 14.
CC FT ZN_FING 534 556 C2H2-type 15.
CC FT ZN_FING 562 584 C2H2-type 16.
CC FT ZN_FING 590 612 C2H2-type 17.
CC FT ZN_FING 618 640 C2H2-type 18.
CC FT ZN_FING 646 668 C2H2-type 19.
CC SQ SEQUENCE 673 AA; 77243 MW; 8C8B1B6D272D08FC CRC64;

Query Match          74.0%; Score 327; DB 1; Length 673;
Best Local Similarity 72.7%; Pred. No. 2.3e-26;
Matches 56; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCGKAFIQKSNLIRHQRTHGKPYKCEGKAFQSSNLTQSKHKKIHTGKPKYCKQC 60
Db 564 CKQCGKAFGSAHLMQHGRTHTGKPYCKQCGKSGFCASRLQMHGRTHGKPKYCKQC 623
QY 61 GKAFGCPNLRHGRTH 77
Db 624 GKAFGCPNLRHGRTH 640

RESULT 3
Q8N8Q4 HUMAN
ID Q8N8Q4_HUMAN PRELIMINARY; PRT; 592 AA.
AC Q8N8Q4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ39023.
GN Name=ZNF505;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Hominidae; Homo.

NCBI_TaxID=9606;

NCBI

NUCLEOTIDE SEQUENCE.

Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Makita H., Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai T., Isono Y., Nakamura Y., Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hosida S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O., Nakamura Y., Togiwa S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Mueshino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshihawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;

"Complete sequencing and characterization of 21,243 full-length human cDNAs";

Nat. Genet. 36:40-45(2004).

EMBL; AK96342; BAC04764.1; -; mRNA.

HSSP; P08047; 1SP2.

Ensembl; ENSG00000081665; Homo sapiens.

HGNC; HGNC:37799; ZNF505.

GO; GO:0005634; C:nucleus; IEA.

GO; GO:0046912; F:metal ion binding; IEA.

GO; GO:0003676; F:nucleic acid binding; IEA.

GO; GO:0008270; F:zinc ion binding; IEA.

GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

GO; GO:0006350; P:transcription; IEA.

InterPro; IPR001909; KRAB.

InterPro; IPR006025; Pept_M_Zn_BS.

InterPro; IPR007087; Znf_C2H2.

InterPro; IPR007086; Znf_C2H2_sub.

Pfam; PF01352; KRAB; 1.

Pfam; PF00096; zf-C2H2; 15.

PRINTS; PR00048; ZINCFINGER.

ProDom; PD000003; Znf_C2H2; 15.

SMART; SM00349; KRAB; 1.

SMART; SM00355; Znf_C2H2; 16.

PROSITE; PS00355; KRAB; 1.

PROSITE; PS00028; ZINC_FINGER_C2H2_1; 15.

PROSITE; PS00157; ZINC_FINGER_C2H2_2; 16.

PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 2.

SEQUENCE 592 AA; 67920 MW; D116P9366EB4BDE CRC64;

Query Match 72.2%; Score 319; DB 2; Length 592;

Best Local Similarity 70.1%; Pred. No. 1.4e-25;

Matches 54; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Qy 1 KCDCGKAFIQKSNLIRHQRTHTGKPKYCECGKAFQSSNLTGKHKIHTGKPKYKQC 60

Db 427 CEECGKAFNQSSSLTKHKIHTGKPKYCECGKAFNQSSSLTKHKIHTGKPKYKCEC 486

Qy 61 GKAFGCPSPNLRHRGRTHT 77

InterPro: IPR007087; Znf_C2H2.
InterPro: IPR007086; Znf_C2H2_sub.
PFam: PF01352; KRAB; 1.
PFam: PF00096; zf-C2H2; 16.
PRINTS: PR00048; ZINCFINGER.
ProDom: PD000003; Znf_C2H2; 16.
SMART: SM00349; KRAB; 1.
SMART: SM00355; Znf_C2H2; 17.
PROSITE: PS0028; ZINC_FINGER_C2H2_1; 16.
PROSITE: PS0028; ZINC_FINGER_C2H2_2; 17.
PROSITE: PS0157; ZINC_FINGER_C2H2_1; 16.
PROSITE: PS0157; ZINC_FINGER_C2H2_2; 17.
Dna-binding: Metal-binding; Nuclear protein; Repeat; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT DOMAIN <1 71
FT ZN FING 141 163 C2H2-type 1; degenerate.
FT ZN FING 169 191 C2H2-type 2.
FT ZN FING 197 219 C2H2-type 3.
FT ZN FING 225 247 C2H2-type 4.
FT ZN FING 253 275 C2H2-type 5.
FT ZN FING 281 303 C2H2-type 6.
FT ZN FING 309 331 C2H2-type 7.
FT ZN FING 337 359 C2H2-type 8.
FT ZN FING 365 387 C2H2-type 9.
FT ZN FING 393 415 C2H2-type 10.
FT ZN FING 421 443 C2H2-type 11.
FT ZN FING 449 471 C2H2-type 12.
FT ZN FING 477 499 C2H2-type 13.
FT ZN FING 505 527 C2H2-type 14.
FT ZN FING 533 555 C2H2-type 15.
FT ZN FING 561 583 C2H2-type 16.
FT ZN FING 589 611 C2H2-type 17.
FT ZN FING 71 71 S -> SGP (in Ref. 2).
FT CONFLICT 85 85 Q -> H (in Ref. 2).
FT CONFLICT 115 115 S -> R (in Ref. 2).
FT CONFLICT 132 132 C -> S (in Ref. 2).
FT CONFLICT 149 149 V -> D (in Ref. 2).
FT NON TER 1
FT SEQUENCE 616 AA; 70572 MW; DF309883AB61160A CRC64;
Query Match 72.2%; Score 319; DB 1; Length 616;
Best Local Similarity 70.1%; Pred. No. 1.5e-25;
Matches 54; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
QY 1 CKDCGKAFQKSNLRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 451 CEECGKAFNQSSSLTKHKIHTGKPYKCECGKAFNQSSSLTKHKIHTGKPYKCEEC 510
QY 61 GKAFGCPNLRHGRTH 77
Db 511 GRAFNQSSLTIKHKIHT 527
RESULT 5
Q9H9X5 HUMAN PRELIMINARY; PRT; 173 AA.
AC Q9H9X5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ12488.
GN Name=ZNF505;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Ngai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
Nagahari K., Murakami K., Yasuda T., Kaku Y., Kodaira H., Kondo H.,
Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
Fujimori K., Tanai H., Kimata M., Watanabe S., Yosida M., Hotuta T.,
Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Okamoto S.,
Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Senba T.,
Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
Togaashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
"Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
RL Nat. Genet. 36:40-45(2004).
DR ENBL; AK022550; BAB14093.1; -; mRNA.
DR HSP; P08047; ISP2.
DR Ensembl; ENSG0000081665; Homo sapiens.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2; 6.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 6.
DR SMART; SM00355; Znf_C2H2; 6.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 173 AA; 19543 MW; 79332B6AC4C26424 CRC64;
Query Match 71.3%; Score 315; DB 2; Length 173;
Best Local Similarity 70.1%; Pred. No. 1.1e-25;
Matches 54; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
QY 1 CKDCGKAFQKSNLRHQRTHTGKPYKCECGKAFQSSNLTQKHKIHTGKPYKCKQC 60
Db 92 CEECGKAFNQSSSLTKHKIHTGKPYKCECGKAFNQSSSLTKHKIHTGKPYKCEEC 151
QY 61 GKAFGCPNLRHGRTH 77
Db 152 GKAFISPSLSRHEIHT 168
RESULT 6
Q95FB0 HUMAN PRELIMINARY; PRT; 403 AA.
ID Q95FB0 HUMAN PRELIMINARY; PRT; 403 AA.
AC Q95FB0;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE PREDICTED: KRAB domain only 2 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Ngai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
Nagahari K., Murakami K., Yasuda T., Kaku Y., Kodaira H., Kondo H.,
Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
Fujimori K., Tanai H., Kimata M., Watanabe S., Yosida M., Hotuta T.,
Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Okamoto S.,
Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Senba T.,
Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
Togaashi T., Oyama M., Hata H.,

SQ SEQUENCE 503 AA; 57886 MW; 67B3EF2978354A2A CRC64;
 Query Match 71.3%; Score 315; DB 2; Length 503;
 Best Local Similarity 68.8%; Pred. No. 3.3e-25;
 Matches 53; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 Qy 1 CKDCGKAFIQKSNLIRHQHTGKEKPKYCEEGKAFQTQSNLTGKHKIHTGKPKYCKQC 60
 Db 420 CEBCGKAFWSAGLHKRRHTGKEKPKYCEEGKAYTTSNLTGKHTGKPKYCKKEC 479
 Qy 61 GKAFGCPSNLRHGRTH 77
 Db 480 GKAFNWSDDLHKHRIH 496
 RESULT 8
 Q8NA64 HUMAN
 ID ID Q8NA64 HUMAN PRELIMINARY; PRT; 530 AA.
 AC Q8NA64;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein FLJ35804.
 GN Name=ZNF568;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 ON NCBI_TaxID=9606;
 RX [1]
 RP NCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX PubMed=14702039; DOI=10.1038/ngl1285;
 RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Horuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togliya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiya H., Satoh N., Takami S., Teraehima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Iton T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isoigai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR EMBL; AK093123; BAC04064.1; -; mRNA.
 DR HSSP; P07248; 1PAA.
 DR Ensembl; ENSG00000198453; Homo sapiens.
 DR HGNC; HGN:25392; ZNF568.
 DR GO; GO:0003634; C:nucleus; IEA.
 DR GO; GO:0046672; F:metal ion binding; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent;
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent;

```
DR GO:0006350; P:transcription; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 13.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 12.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 13.
FT NON_TER 530 530
SQ SEQUENCE 530 AA; 61275 MW; 9491D3767BF5C5E4 CRC64;

Query Match 71.3%; Score 315; DB 2; Length 530;
Best Local Similarity 71.4%; Pred. No. 3.4e-25;
Matches 55; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 CKDCGKAPIQKSNLIRHQRTHGTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCKQC 60
Db 412 CSECGKAPIQKSNLIRHQRTHGTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCKQC 471
QY 61 GKAFGCPNLRHGRTH 77
Db 472 GKAFSQRNLEHEKIH 488

RESULT 9
Q6ZMR0 HUMAN
ID Q6ZMR0_HUMAN PRELIMINARY; PRT; 563 AA.
AC Q6ZMR0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16755.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cerebellum;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Negahari K., Maehuo Y., Nagai K., Isegai T.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK131527; BAD18665.1; -; mRNA.
DR HSP; P08046; 1A1G.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 14.
DR ProDom; PD000003; Znf_C2H2; 14.
DR PRINTS; PR00048; ZINCFINGER.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 14.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.

DR GO:000157; ZINC_FINGER_C2H2_2; 15.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 563 AA; 64779 MW; 99792DD8927BFC2A CRC64;

Query Match 71.3%; Score 315; DB 2; Length 563;
Best Local Similarity 71.4%; Pred. No. 3.7e-25;
Matches 55; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCGKAPIQKSNLIRHQRTHGTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCKQC 60
Db 254 CEECGKAFNRSSNLTGKIVHTGKPYKCEGKAFQSSNLTGKKIHTGKPYKCGEC 313
QY 61 GKAFGCPNLRHGRTH 77
Db 314 GKFTLSHLTHKRIH 330

RESULT 10
ZNF91 HUMAN
ID ZNF91_HUMAN STANDARD; PRT; 1191 AA.
AC Q05481;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zinc finger protein 91 (Zinc finger protein HTF10) (HPF7).
GN Name=ZNF91;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9323677; PubMed=8467795;
RA Bellefroid E.J., Marine J.C., Ried T., Lecocq P.J., Riviere M.,
RA Ameiya C.T., Poncelet D.A., Coulie P.G., de Jong P.J., Szpirer C.,
RA Ward D.C., Martial J.A.;
RT "Clustered organization of homologous KRAB zinc-finger genes with
RT enhanced expression in human T lymphoid cells.";
RL EMBL J. 12:1363-1374 (1993).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 15-204.
RX MEDLINE=91219421; PubMed=2023909;
RA Bellefroid E.J., Poncelet D.A., Lecocq P.J., Revelant O.,
RA Martial J.A.;
RT "The evolutionarily conserved Kruppel-associated box domain defines a
RT subfamily of eukaryotic multifingered proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3608-3612 (1991).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
CC family.
CC -!- SIMILARITY: Contains 36 C2H2-type zinc fingers.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC -!- CAUTION: The sequence from position 1159 to the C-terminal is
CC derived from the translation of an Alu repeat.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; L11672; AAA59469.1; -; mRNA.
DR EMBL; M61871; AAA58672.1; ALT_SEQ; mRNA.
DR PIR; S35305; S35305.
DR HSSP; P08047; 13P2.
DR SMR; Q05481; 291-372.
DR Ensembl; ENSG00000167232; Homo sapiens.
DR HGNC; HGNC:13166; ZNF91.
DR MIM; 603971; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0008270; F:zinc ion binding; NAS.
```


GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; zf-C2H2; 32.
 DR PRINTS; PR00048; ZINCFINGER.
 DR ProDom; PD000003; Znf_C2H2; 33.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 34.
 DR PROSITE; PS50805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 31.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 35.
 KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Transcription;
 Transcription regulation; Zinc; Zinc-finger.
 FT DOMAIN 13 84
 FT ZN_FING 154 176 C2H2-type 1; degenerate.
 FT ZN_FING 179 200 C2H2-type 2.
 FT ZN_FING 208 232 C2H2-type 3.
 FT ZN_FING 238 260 C2H2-type 4.
 FT ZN_FING 266 288 C2H2-type 5.
 FT ZN_FING 294 316 C2H2-type 6.
 FT ZN_FING 322 344 C2H2-type 7.
 FT ZN_FING 350 372 C2H2-type 8.
 FT ZN_FING 378 400 C2H2-type 9.
 FT ZN_FING 406 428 C2H2-type 10.
 FT ZN_FING 434 456 C2H2-type 11.
 FT ZN_FING 462 484 C2H2-type 12.
 FT ZN_FING 490 512 C2H2-type 13.
 FT ZN_FING 518 540 C2H2-type 14; degenerate.
 FT ZN_FING 546 568 C2H2-type 15.
 FT ZN_FING 574 596 C2H2-type 16.
 FT ZN_FING 602 624 C2H2-type 17.
 FT ZN_FING 630 652 C2H2-type 18.
 FT ZN_FING 658 680 C2H2-type 19.
 FT ZN_FING 686 708 C2H2-type 20.
 FT ZN_FING 714 736 C2H2-type 21.
 FT ZN_FING 742 764 C2H2-type 22.
 FT ZN_FING 770 792 C2H2-type 23.
 FT ZN_FING 798 820 C2H2-type 24.
 FT ZN_FING 826 848 C2H2-type 25.
 FT ZN_FING 854 876 C2H2-type 26.
 FT ZN_FING 885 904 C2H2-type 27; degenerate.
 FT ZN_FING 910 932 C2H2-type 28.
 FT ZN_FING 938 960 C2H2-type 29.
 FT ZN_FING 966 988 C2H2-type 30.
 FT ZN_FING 994 1016 C2H2-type 31.
 FT ZN_FING 1022 1044 C2H2-type 32.
 FT ZN_FING 1050 1072 C2H2-type 33.
 FT ZN_FING 1078 1100 C2H2-type 34.
 FT ZN_FING 1106 1128 C2H2-type 35.
 FT ZN_FING 1134 1156 C2H2-type 36.
 SQ SEQUENCE 1191 AA; 137226 MW; 581056B811B9716D CRC64;
 Query Match 71.3%; Score 315; DB 1; Length 1191;
 Best Local Similarity 72.7%; Pred. No. 8e-25;
 Matches 56; Conservative 6; Mismatches 15; Indels 0; Gaps 0;
 QY 1 CKDCGKAFQKLNLRHQHTGKPKCEGCKAFTQSSNLTGKHKIHTGKPKCKQC 60
 DB 772 KCECGKAFWSLTTRHKRIHTGKPKCEGCKAFSRSTLTGKHKIHTGKPKCKEC 831
 QY 61 GKAFGCPNLRBGRTH 77
 DB 832 GKAFKXSSALAKHKI 848
 RESULT 11
 ID_ZNF_HUMAN STANDARD; PRT; 292 AA.
 AC Q15072;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Zinc finger protein OZF.
 GN Name=ZNF146; Synonyms=OZF;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mammary gland;
 RX MEDLINE=94149744; PubMed=8107129;
 RA le Chalony C., Prosperi M.-T., Haluza R., Apiou F., Dutrillaux B.,
 RA Goubin G.;
 RT "The OZF gene encodes a protein consisting essentially of zinc finger
 RT motifs.";
 RL J. Mol. Biol. 236:399-404(1994).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21203548; PubMed=11306801;
 RA Pibouin L., Villaudy J., Prosperi M., Goubin G.;
 RT "Genomic organization and promoter identification of ZNF146, a gene
 RT encoding a protein consisting solely of zinc finger domains.";
 RL Eur. J. Biochem. 236:991-995(1996).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Liver, skeletal and heart muscle, mammary
 CC cells. Very low levels in brain, lung, placenta and kidney.
 CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
 CC family.
 CC -!- SIMILARITY: Contains 10 C2H2-type zinc fingers.
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/OZFID267.html".
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; X70394; CAA49844.1; -; mRNA.
 CC EMBL; AJ011806; CAB41967.1; -; Genomic_DNA.
 CC PIR; S43826; S43826.
 CC HSP; P08047; ISP2.
 CC TRANSPAC; T02323; -.
 CC DR Ensembl; ENSG00000167635; Homo sapiens.
 CC HGNC; HGNC:12931; ZNF146.
 CC MIM; 601505; -.
 CC GO; GO:0005634; C:nucleus; TAS.
 CC GO; GO:0003677; F:DNA binding; TAS.
 CC GO; GO:0008201; F:heparin binding; TAS.
 CC GO; GO:0008270; F:zinc ion binding; TAS.
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
 CC InterPro; IPR007087; Znf_C2H2.
 CC InterPro; IPR007086; Znf_C2H2_sub.
 CC Pfam; PF00096; zf-C2H2; 10.
 CC PRINTS; PR00048; ZINCFINGER.
 CC ProDom; PD000003; Znf_C2H2; 10.
 CC SMART; SM00355; Znf_C2H2; 10.
 CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 10.
 CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 10.
 KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Zinc;
 KW Zinc-finger.
 FT ZN_FING 16 38 C2H2-type 1.
 FT ZN_FING 44 66 C2H2-type 2.
 FT ZN_FING 72 94 C2H2-type 3.

```

RESULT 13
Q5RFP4_PONPY Q5RFP4_PONPY PRELIMINARY; PRT; 292 AA.
AC Q5RFP4;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp469K1922.
GN Name=DKFZp469K1922;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Pongo.
NCBI_TaxID=9600;
[1]
RN
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG The German cDNA Consortium;
RA Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR857109; CAH89413.1; -; mRNA.

```

[illegible]

GN name=ZNR502;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 [1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP TISSUE=Testis;
 RC MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May be involved in transcriptional regulation.
 CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
 CC -!- SIMILARITY: Belongs to the krueppel C2H2-type zinc-finger protein
 CC family.
 CC -!- SIMILARITY: Contains 14 C2H2-type zinc fingers.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 DR EMBL; BC028377; AAH28377.1; -; mRNA.
 DR HSSP; Q87B25; 348-429.
 DR SMR; Q87B25; 348-429.
 DR Ensembl; ENSG00000196653; Homo sapiens.
 DR HGNC; HGNC:23718; ZNF502.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF00996; zf-C2H2; 14.
 DR PRINTS; PR00048; ZINCFINGER.
 DR ProDom; PD000003; Znf_C2H2; 14.
 DR SMART; SM00355; Znf_C2H2; 14.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 14.
 KW DNA-binding; Metal-binding; Nuclear protein; Repeat; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 FT ZN_FING 155 177 C2H2-type 1.
 FT ZN_FING 183 205 C2H2-type 2.
 FT ZN_FING 211 233 C2H2-type 3.
 FT ZN_FING 239 261 C2H2-type 4.
 FT ZN_FING 267 289 C2H2-type 5.
 FT ZN_FING 295 317 C2H2-type 6.
 FT ZN_FING 323 345 C2H2-type 7.
 FT ZN_FING 351 373 C2H2-type 8.
 FT ZN_FING 379 401 C2H2-type 9.
 FT ZN_FING 407 429 C2H2-type 10.
 FT ZN_FING 435 457 C2H2-type 11.
 FT ZN_FING 463 485 C2H2-type 12.
 FT ZN_FING 491 513 C2H2-type 13.
 FT ZN_FING 519 541 C2H2-type 14.
 SQ SEQUENCE 544 AA; 62920 MW; 63437B670A6B22D0 CRC64;

Query Match 71.0%; Score 314; DB 1; Length 544;
 Best Local Similarity 70.1%; Pred. No. 4.5e-25;
 Matches 54; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 CKDCKGKAFIOKSNLRHQRTHGTGKPYKCECGKAPTQSSNLTGKHKHTGKPYKCKQC 60
 DB 409 CSECGKAFIOKSNLRHQRTHGTGKPYKCECGKAPTQSSNLTGKHKHTGKPYKCKQC 468
 QY 61 GKAFGCPNLRHGRTH 77
 DB 469 GKAFHSSSLTEHHRTH 485

RESULT 15

Q6AYV1 RAT PRELIMINARY; PRT; 589 AA.
 AC Q6AYV1_1
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Zinc finger protein 386 (Krueppel-like).
 GN Name=Znf386;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Schmutz J., Myers R.M.,
 RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RA Director MGC Project;
 RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC078841; AAH78841.1; -; mRNA.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00996; zf-C2H2; 9.
 DR PRINTS; PR00048; ZINCFINGER.
 DR ProDom; PD000003; Znf_C2H2; 9.
 DR SMART; SM00349; KRAB_1.
 DR SMART; SM00355; Znf_C2H2; 9.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 9.
 SQ SEQUENCE 589 AA; 68356 MW; 3B0706382972C903 CRC64;

Query Match 71.0%; Score 314; DB 2; Length 589;
 Best Local Similarity 66.2%; Pred. No. 4.9e-25;
 Matches 51; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 CKDCKGKAFIOKSNLRHQRTHGTGKPYKCECGKAPTQSSNLTGKHKHTGKPYKCKQC 60
 DB 456 CAECCKSFTQNSTLSQHORHTGKPYKCECGKAPTQSSNLTGKHKHTGKPYKCKQC 515
 QY 61 GKAFGCPNLRHGRTH 77

Db 516 GRAFNCRSSFTKKRIH 532

Search completed: January 9, 2006, 20:08:48
Job time : 162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: January 11, 2006, 20:01:04 ; Search time 3642 Seconds
(without alignments)
1201.797 Million cell updates/sec

Title: US-10-669-861-2 COPY 33 109

Perfect score: 442
Sequence: 1 CKDCGKAFIQKSNLIRHRT.....KCGKAFCGCPSNLRHGRTH 77

Scoring table: BLOSUM62

```

    scoring cware.  DRO30M02
    Xgapop 10.0 , Xgapext 0.5

```

lgapop	10.0	,	lgapext	0.5
Fgapop	6.0	,	Fgapext	7.0
Delop	6.0	,	Delext	7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frameat_p2n_model -DEV=xlh
-Q/cn2_1/USFTO_spool/US10669861/runat_09012006_144453_20373/app_query.fasta_1.263
-DB=GenEmul -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pt6 -NORM=ex -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10669861@cgn_1_1_4939 @runat_09012006_144453_20373 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEBQUERY -NEG_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

Database :	GenEmbl :
1:	gb_bac
2:	gb_in
3:	gb_en
4:	gb_on
5:	gb_ov
6:	gb_pae
7:	gb_ph
8:	gb_pr
9:	gb_rc
10:	gb_s
11:	gb_s
12:	gb_u
13:	gb_v
14:	gb_h
15:	gb_p

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	329	74.4	2562	6	CQ728633
2	329	74.4	158404	8	AC092835
3	329	74.4	294915	14	AC155005

C	4	327	74.0	672	6	AX721061	Sequence
	5	327	74.0	2192	6	AX540402	Sequence
	6	327	74.0	2275	6	AX468105	Sequence
	7	327	74.0	2280	6	AX835301	Sequence
	8	327	74.0	2280	8	AK098300	Homo sap.
	9	327	74.0	2323	8	BC047412	Homo sap.
C	10	327	74.0	178184	8	AC008770	Homo sap.
	11	327	74.0	179581	14	AC009397	Homo sap.
C	12	327	74.0	180510	8	AC022415	Homo sap.
	13	325	73.5	3300	6	AX022430	Sequence
C	14	322	72.9	216979	14	AC103247	Rattus r.
C	15	322	72.9	238307	14	AC133610	Rattus r.
	16	321	72.6	330	6	AX020476	Sequence
	17	321	72.6	3300	6	AX202426	Sequence
	18	321	72.6	3300	6	AX202427	Sequence
	19	319	72.2	2867	6	AX834102	Sequence
	20	319	72.2	2867	8	AK096342	Sequence
	21	319	72.2	109478	8	AC007204	Homo sap.
	22	318	71.9	197553	14	AC146103	Pan trogl.
	23	316	71.5	2253	6	CQ779570	Sequence
	24	316	71.5	127811	14	AC008375	Homo sap.
C	25	316	71.5	138491	9	AL929042	Mouse DNI
C	26	316	71.5	146877	14	AC028487	Homo sap.
C	27	316	71.5	169602	14	AC079567	Mus musc.
	28	316	71.5	202768	14	AC092345	Homo sap.
	29	316	71.5	215962	8	AC010615	Homo sap.
	30	316	71.5	239241	14	AC128484	Rattus r.
C	31	316	71.5	264661	14	AC094217	Rattus r.
C	32	315	71.3	827	6	BD147256	Primer f.
	33	315	71.3	827	6	AX867194	Sequence
	34	315	71.3	882	6	CQ734128	Sequence
	35	315	71.3	1449	6	BD157779	Primer f.
	36	315	71.3	1449	8	AX879332	Sequence
	37	315	71.3	1449	8	AC022550	Homo sap.
	38	315	71.3	1512	8	BC027608	Homo sap.
	39	315	71.3	1914	6	AX747939	Sequence
	40	315	71.3	1914	8	AK093123	Homo sap.
	41	315	71.3	1997	6	CQ726025	Sequence
	42	315	71.3	2848	8	HSMB07672	Homo sap.
	43	315	71.3	2864	6	AX833387	Sequence
	44	315	71.3	2664	8	AK095062	Homo sap.
	45	315	71.3	3096	6	CQ731632	Sequence

ALIGNMENTS

```

RESULT 1
CQ728633
LOCUS       2562 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION  Sequence 14567 from Patent WO02068579.
ACCESSION   CQ728633
VERSION     CQ728633.1  GI:42297975
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE   1
  AUTHORS   Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 14567 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES    Location/Qualifiers
             1..2562
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN
Alignment Scores:

```

```
Pred. No.: 1.89e-27 Length: 2562
Score: 329.00 Matches: 56
Percent Similarity: 85.71% Conservative: 10
Best Local Similarity: 72.73% Mismatches: 11
Query Match: 74.43% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x CQ728633 (1-2562)

QY 1 CysLysAspCysGlyLysAlaPheLleGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
   |||||
Db 2248 TGCAATGATTGTCACAAAGCCCTTACGTACCGCTTATTCGTATCATCAGAGACA 2307
   |||||

QY 21 HisThrGlyGluLysProTyLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
   |||||
Db 2308 CACACTGGAGAGAACCTTACAAATGTATAGAGTTGTGAAAGACATTACGCCAGACTCA 2367
   |||||

QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyLysCysLysGlnCys 60
   |||||
Db 2368 TCTCTTACAAAGCATCAGAAACTCACACTGGAGAAAGCCCTACAAAGTGAAGGAATGT 2427
   |||||

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
   |||||
Db 2428 GGAAAGCCCTTAGCCAGAGTTTATCCCTCTCTCAACATCAGAAACTCAT 2478
   |||||

RESULT 2
AC092835
LOCUS AC092835 158404 bp DNA linear PRI 15-APR-2005
DEFINITION Homo sapiens BAC clone RP11-11P22 from 2, complete sequence.
ACCESSION AC092835
VERSION AC092835.4 GI:19071676
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 158404)
AUTHORS Doebber,A., Nguyen,C. and Haglund,K.
TITLE The sequence of Homo sapiens BAC clone RP11-11P22
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 158404)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 3 (bases 1 to 158404)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 158404)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 158404)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (15-APR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Mar 2, 2002 this sequence version replaced gi:18250106.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0011P22
-----
```

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-708D7; the clone sequenced to the right is RP11-468G5, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-11P22.

There is a transposon that has been omitted from the submitted sequence. The transposons should insert after base position 141089. There is a PCR only region from 145576 to 145630.

Polymorphisms have been identified between AC093621 and AC092835. Data from AC093621, AC103563 and AC026626 was used to finish this clone, AC092835.

FEATURES

```
1. 158404
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-11P22"
/clone_lib="RPCI-11"
/complement(5713..40509)
/gene="MRPS5"
complement(join(5713..6086,8891..9027,18979..19041,
19342..19399,20182..20228,23145..23235,24928..24962,
26680..26913,28421..28546,33571..33708,36370..36450,
40239..40509))
/gene="MRPS5"
complement(join(5862..6086,8891..9027,18979..19041,
19342..19399,20182..20228,23145..23235,24928..24962,
26680..26913,28421..28546,33571..33708,36370..36450,
40239..40296))
/gene="MRPS5"
/note="Homo sapiens mitochondrial ribosomal protein S5
(MRPS5), nuclear gene encoding mitochondrial protein,
mRNA; H_NH0011P22.1
This gene was based on gi(16554614)"
/codon_start=1
/product="unknown"
/protein_id="AAK88980.1"
/db_xref="GI:62630235"
/translation="MATAVRAVGCLPVLCSGTAGHLGROCSLNTLPASILAWKSVL
GNHLSLIGTRDTHPYASLSRALQCCISPSHLMQVQRPYSFETKLTADLWKGA
LAETGAGAKGRGRTKKKKKDLNRQGIIGEGYGLFPLGLNPLMKNQAVQITQR
SKEQEKVADMTIQREWRKKMKVGRGWSNGWSGLSPDPGPGCQETDF
DTRILRVNFTTAKERKKSIRLVAVNGKGAAGFSIGKATDRMDAFRAKNRVA
HHLHYIERYEDTIFHDIILRFKRTTHIMKKQPKGYGLRCHRAITICRLGIDMYA
KVSGSINWLSLTQGLFRLSRQETHQQLADKGLHVVIREECGCLPIVWASPRGLR
```

misc_feature	KDPEDEVDVVDKLDHEDVKTAAQMKRKSVMNSWNLKRAAT"		
	8404..8639		
	/note="CpG_island (%GC=61.4, o/e=0.95, #CpGs=20)"		
	39850..40711		
	/note="CpG_island (%GC=69.5, o/e=0.85, #CpGs=92)"		
gene	complement(66160..78023)		
	/gene="FLJ14457"		
mRNA	complement(join(66160..68772,71179..71274,71638..71764,75856..75944,77745..78023))		
	/gene="FLJ14457"		
CDS	complement(join(67787..68772,71179..71274,71638..71758))		
	/note="Homo sapiens zinc finger protein 514 (ZNF514), mRNA.; H_NH0011P22.2		
	This gene was based on gi(14249455)"		
	/codon_start=1		
	/product="unknown"		
misc_feature	/protein_id="AAK8981.1"		
	/db_xref="GI:62630236"		
	/translation="MTFEDVAVFQSWNGQLNPAQKDYREVMLENFRNLAILGLLV SKPYVLCQEEGEPFVREISTGAHDKRRKSKSPGWSGISKBELFQVVSVEK HIQDVLQSKLAACDCGQLEMQIQKQERHLKQMSITIHKSATTLSDRYKMGFGRSL GLRVLVNOHSILMGSGYKCDPTEPQTGLGNNSTHPEKSKCKNCGKSPHQSE LARHORCHTGEKPYECSDCGRAFGLHSLIKQRTHTGKPYECSCGGRAPQSSSLV LHYRFTGKPYKNCBGRAFGLHSLIKQRTHTGKPYECRECGRTFSQSSSLIVH YRFTGKPYKNCBGRAFSPQSSSLQHYRFTHTGKPYKNCBGRAFHTASLTKHQ R SHAKKTL"		
	77563..78481		
	/note="CpG_island (%GC=76.1, o/e=0.79, #CpGs=108)"		
misc_feature	83918..84593		
	/note="CpG_island (%GC=61.5, o/e=0.98, #CpGs=57)"		
gene	83950..101481		
	/gene="ZNF2"		
mRNA	join(83950..84365,94693..94764,95988..96114,98664..98777,99608..101481)		
	/gene="ZNF2"		
CDS	99733..100614		
	/gene="ZNF2"		
	/note="Homo sapiens zinc finger protein 2 (A1-5) (ZNF2), mRNA.; H_NH0011P22.3		
	This gene was based on gi(20304090)"		
	/codon_start=1		
misc_feature	/product="unknown"		
	/protein_id="AAK8982.1"		
	/db_xref="GI:62630237"		
	/translation="MTEKQSPGETRKKSLRDKGLRPRRSLGREILTKERHOECS DCGKTFPDSHSLTRHRTHTGKPYDCRCGKAFHSRSLRHLSHGTSYECVC SKAFDRSLTRHRTHTGKPYDCRCGKAFHSRSLRHLSHGTSYECVCHQCGK AFQSKSLTRHRTHTGKPYDCRCGKAFYGVSSLRHOKAHADGPRYQNCCKGKAF FDRSLTRHRTHTGKPYDCRCGKAFYGVSSLRHOKAHADGPRYQNCCKGKAF KSSVIQHORRYAKQID"		
	102198..102424		
misc_feature	/note="CpG_island (%GC=58.6, o/e=0.94, #CpGs=17)"		
	125652..126308		
misc_feature	/note="CpG_island (%GC=70.6, o/e=0.87, #CpGs=68)"		
	151899..152168		
misc_feature	/note="CpG_island (%GC=60.7, o/e=0.74, #CpGs=20)"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.9e-25	Length:	158404
Score:	329.00	Matches:	56
Percent Similarity:	85.71%	Conservative:	10
Best Local Similarity:	72.73%	Mismatches:	11
Query Match:	74.43%	Indels:	0
DB:	8	Gaps:	0
US-10-669-861-2_COPY_33_109 (1-77) x AC092835 (1-158404)			
QY	1 CysLysAspCysGlyLysAlaPheHleGlnLysSerAsnLeuLeuArgHisGlnArgThr 20		
Db	133797 TGCATGATGTTGGCCAAAGCCCTTCAAGTACCGCTGAGCCCTTATTCGTATCAGAGAAC 133856		

QY	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db	133857	CACACTGGAGAGAAACCTTACAAATGTAAGATTGTGGAAGACCATTCAGCCAGACTCA 133916
QY	41	AsnLeuThrLysHisLysLysLeHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db	133917	TCTCTTACAAAGCATCAGAAACTTACACTGGAGAAAGCCCTTACAGTGTAGGAATGT 133976
QY	61	GlyLysAlaPheGlyCysProSerAsnLeuArgGArgHisGlyArgThrHis 77
Db	133977	GGAAAGCCCTTACAGCAGAGTTCCTCTCTCAACATCAGAAAAAATCAT 134027
RESULT 3		
LOCUS	AC155005	294915 bp DNA linear HTG 01-JUL-2005
DEFINITION	Bos taurus clone CH240-36J15, *** SEQUENCING IN PROGRESS ***, 48	
ACCESSION	AC155005	
VERSION	AC155005.3 GI:68265223	
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.	
SOURCE	Bos taurus (cow)	
ORGANISM	Bos taurus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.	
	Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,N., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munitasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K., Nwaokemeleho,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von	

Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 294915)
 Worley, K.C.
 Direct Submission
 Submitted (07-JAN-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 294915)
 Cow Genome Sequencing Consortium.
 Direct Submission
 Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 28, 2005 this sequence version replaced gi:57334887.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: FCCE
 Center clone name: CH240-36J15
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 262862 bases at least Q40
 Consensus quality: 266947 bases at least Q30
 Consensus quality: 270751 bases at least Q20
 Estimated insert size: 269122; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one 'clone'
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 48 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2521: contig of 2521 bp in length
 * 2522 2571: gap of 50 bp
 * 2572 5531: contig of 2960 bp in length
 * 5532 5631: gap of unknown length
 * 5632 7766: contig of 2135 bp in length
 * 7767 7816: gap of 50 bp
 * 7817 15053: contig of 7237 bp in length
 * 15054 15103: gap of 50 bp
 * 15104 19441: contig of 4338 bp in length
 * 19442 19491: gap of 50 bp
 * 19492 23159: contig of 3668 bp in length
 * 23160 23259: gap of unknown length
 * 23260 24997: contig of 1738 bp in length
 * 24998 25047: gap of 50 bp
 * 25048 45806: contig of 20759 bp in length
 * 45807 45856: gap of 50 bp
 * 45857 53466: contig of 7610 bp in length

* 53467 53516: gap of 50 bp
 * 53517 58591: contig of 5075 bp in length
 * 58592 58641: gap of 50 bp
 * 58642 66086: contig of 7445 bp in length
 * 66087 66136: gap of 50 bp
 * 66137 81849: contig of 15713 bp in length
 * 81850 81899: gap of 50 bp
 * 81900 84423: contig of 2524 bp in length
 * 84424 96159: gap of 11736 bp
 * 96160 100621: contig of 4462 bp in length
 * 100622 100671: gap of 50 bp
 * 100672 104944: contig of 4273 bp in length
 * 104945 104995: gap of 50 bp
 * 104996 112201: contig of 7207 bp in length
 * 112202 112251: gap of 50 bp
 * 112252 132362: contig of 20111 bp in length
 * 132363 132412: gap of 50 bp
 * 132413 140044: contig of 7632 bp in length
 * 140045 140094: gap of 50 bp
 * 140095 162586: contig of 22492 bp in length
 * 162587 162636: gap of 50 bp
 * 162637 170570: contig of 7934 bp in length
 * 170571 171490: gap of 920 bp
 * 171491 175896: contig of 4406 bp in length
 * 175897 175946: gap of 50 bp
 * 175947 183357: contig of 5411 bp in length
 * 183358 183373: gap of 1016 bp
 * 183374 183788: contig of 1415 bp in length
 * 183789 183988: gap of 200 bp
 * 183989 212516: contig of 28528 bp in length
 * 212517 214190: gap of 1674 bp
 * 214191 216319: contig of 2129 bp in length
 * 216320 216419: gap of unknown length
 * 216420 222050: contig of 5631 bp in length
 * 222051 222100: gap of 50 bp
 * 222101 226487: contig of 4387 bp in length
 * 226488 226587: gap of unknown length
 * 226588 228131: contig of 1544 bp in length
 * 228132 228365: gap of 234 bp
 * 228366 231899: contig of 3534 bp in length
 * 231900 231999: gap of unknown length
 * 232000 233871: contig of 1872 bp in length
 * 233872 233921: gap of 50 bp
 * 233922 240058: contig of 6137 bp in length
 * 240059 240108: gap of 50 bp
 * 240109 254341: contig of 14233 bp in length
 * 254342 254441: gap of unknown length
 * 254442 259967: contig of 5526 bp in length
 * 259968 260017: gap of 50 bp
 * 260018 263868: contig of 3851 bp in length
 * 263869 263968: gap of unknown length
 * 263969 265021: contig of 1053 bp in length
 * 265022 265121: gap of unknown length
 * 265122 266132: contig of 1011 bp in length
 * 266133 266232: gap of unknown length
 * 266233 267543: contig of 1311 bp in length
 * 267544 267643: gap of unknown length
 * 267644 268658: contig of 1015 bp in length
 * 268659 268758: gap of unknown length
 * 268759 270271: contig of 1513 bp in length
 * 270272 270371: gap of unknown length
 * 270372 271638: contig of 1267 bp in length
 * 271639 271738: gap of unknown length
 * 271739 273324: contig of 1586 bp in length
 * 273325 273424: gap of unknown length
 * 273425 274728: contig of 1304 bp in length
 * 274729 274828: gap of unknown length
 * 274829 276302: contig of 1474 bp in length

Alignment Scores: 3.82e-25 Length: 294915
 Pred. No.: 329.00 Matches: 56
 Score: 85.71% Conservative: 10
 Percent Similarity:


```
Best Local Similarity: 72.73% Mismatches: 11
Query Match: 74.43% Indels: 0
DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC155005 (1-294915)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 44564 TGCACGATGTGCCAAAGCCTTCAGTGACCGTTCAGCCCTTATCCGTATCAGAGAAC 44623

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 44624 CACACTGGAGAGAAACCTTACAAAGTGAAGCACTGGGAAAGCCTTCAGCCAGACTCA 44683

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 44684 TCTCTTACAAGCATCAGAAACTCACCTGGGAGAAACCCCTATTAAGTGAAGCAATGT 44743

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuAArgHisGlyArgThrHis 77
Db 44744 GGAAAGCCTTCAGCCAGAGTTCATCCCTTCTCAACATCAGAAACTCAT 44794

RESULT 4
LOCUS AX721061 672 bp mRNA linear PAT 07-MAY-2003
DEFINITION Sequence 21 from Patent WO0220754.
ACCESSION AX721061
VERSION AX721061.1 GI:30421897
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Hominidae; Homo.
REFERENCE
AUTHORS Lincoln,S.E., Altus,C.M., Dufour,G.E., Chalup,M.S., Hillman,J.L.,
Jones,A., Yu,J.Y., Wright,R.J., Gietzen,D., Liu,Toomy,F., Yap,P.,
Dahl,C.R., Momiyama,M., Bradley,D., Rohatgi,S., Harris,B.,
Roseberry,A.M., Gerstin,E.H., Peralta,C.H., David,M., Panzer,S.,
Flores,V., Daffo,A., Marwaha,R., Chen,A., Chang,S.C. and Inman,R.R.
Molecules for diagnostics and therapeutics
Patent: WO 0220754-A 21 14-MAR-2002;
Incye Genomics, Inc. (US)
FEATURES
source
1..672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/note="Incye ID No: LG:1080545.1:2000SEP08"

ORIGIN
Alignment Scores:
Pred. No.: 7.15e-28 Length: 672
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AX721061 (1-672)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 311 TGTAGCAGTGTGGGAAAGCCTTTGGATCGCTCACACCTTCAATGATGATGAGGACT 370

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 371 CACACTGGAGAGAAACCCCTATCAATGTAAGCAGTGTGGGAACTCTTTGGATGCTCCTCG 430

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 431 CGACTTCAATGATGGAAGGACTCACACTGGAGAGAAACCGTATATAATGTAAGCAATGT 490
```

```
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuAArgHisGlyArgThrHis 77
Db 491 GGGAAAGCCTTTGGATGTCCTCAAACTTCGAGGCATGGAAGGACTCAC 541

RESULT 5
LOCUS AX540402 2192 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 14 from Patent WO02055738.
ACCESSION AX540402
VERSION AX540402.1 GI:25273419
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Hominidae; Homo.
REFERENCE
AUTHORS Panzer,S.R., Lincoln,S.E., Altus,C.M., Dufour,G.E., Hillman,J.L.,
Jones,A.L., Dam,T.C., Liu,T.F., Harris,B., Flores,V., Daffo,A.,
Marwaha,R., Chen,A.J., Chang,S.C., Gerstin,E.H., Peralta,C.H.,
David,M.H. and Lewis,S.A.
Molecules for disease detection and treatment
Patent: WO 02055738-A 14 18-JUL-2002;
Incye Genomics, Inc. (US)
FEATURES
source
1..2192
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incye ID No: LI:1175131.1:2001JAN12"

ORIGIN
Alignment Scores:
Pred. No.: 2.68e-27 Length: 2192
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AX540402 (1-2192)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 1845 TGTAGCAGTGTGGGAAAGCCTTTGGATCGCTCACACCTTCAATGATGAGGACT 1904

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 1905 CACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAACTCTTTGGATGCTCCTCG 1964

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1965 CGACTTCAATGATGGAAGGACTCACACTGGAGAGAAACCGTATATAATGTAAGCAATGT 2024

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuAArgHisGlyArgThrHis 77
Db 2025 GGGAAAGCCTTTGGATGTCCTCAAACTTCGAGGCATGGAAGGACTCAC 2075

RESULT 6
LOCUS AX468105 2275 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 28 from Patent WO0250279.
ACCESSION AX468105
VERSION AX468105.1 GI:21900979
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Hominidae; Homo.
REFERENCE
AUTHORS Baughin,M.R., Lu,Y., Arvizu,C., Ramkumar,J., Yao,M.G.,
Policky,J.L., Wallia,N.K., Tribouley,K.M., Yue,H., Batra,S.,
```

Ding,L., Lal,P.G., Borowsky,M.L., Lu,D.A., Gandhi,A.R., Griffin,J.A., Xu,Y., Azimzai,Y., Gietzen,K.J., Tang,Y.T., Warren,B.A., Mason,P.M., Burford,N., Hafalia,A.J., Lee,E.A., Yang,J., Gorvad,A.E., Emerling,B.M., Marquis,J.P., Lee,S.Y., Swarnakar,A. and Reddy,R.
Nucleic acid-associated proteins
Patent: WO 0250279-A 28 27-JUN-2002;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1..2275
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 5435937CB1"

ORIGIN

Alignment Scores:
Pred. No.: 2.8e-27 Length: 2275
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AX468105 (1-2275)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 1914 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAATGCATGGAAGGACT 1973
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1974 CACACTGGAGAGAACCTTATGAATGTAGCAGTGTGGAGTCTTTGGATGTCCTCG 2033
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 2034 CGACTTCAAATGCATGGAAGGACTCACACTGGAGAGAACCGTATAAATGTAAGCAATGT 2093
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 2094 GGGAAAGCTTTTGGATGTCTCCCTCAAACCTTCGAAGGCATGGAAGGACTCAC 2144

RESULT 7

AX835301 LOCUS 2280 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 2425 from Patent EP1347046.
ACCESSION AX835301
VERSION AX835301.1 GI:39921436
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Masuho,Y.

TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 2425 24-SEP-2003;
Research Association for Biotechnology (JP)

FEATURES

source
1..2280
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2.8e-27 Length: 2280
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7

Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x AX835301 (1-2280)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 1861 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAATGCATGGAAGGACT 1920
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1921 CACACTGGAGAGAACCTTATGAATGTAGCAGTGTGGAGTCTTTGGATGTCCTCG 1980
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 1981 CGACTTCAAATGCATGGAAGGACTCACACTGGAGAGAACCGTATAAATGTAAGCAATGT 2040
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 2041 GGGAAAGCTTTTGGATGTCTCCCTCAAACCTTCGAAGGCATGGAAGGACTCAC 2091

RESULT 8

AX098300 LOCUS 2280 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40981 fig, clone UTERU2014548, moderately similar to Homo sapiens ZKI mRNA for Kruppel-type zinc finger protein.
ACCESSION AK098300
VERSION AK098300.1 GI:21758285
KEYWORDS oligo capping; fig (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2280)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

source
1..2280
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2014548"
/tissue_type="uterus"
/clone_lib="UTERU2"
/note="Cloning vector: pME18SFL3"
172..2193
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC05279.1"

CDS

US-10-669-861-2 COPY 33 109 (1-77) x BC047412 (1-2323)

Qy	1	CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr	20
		::::	
Db	1865	TGTAAGCAGTGTGGGAAGCGCTTTGGATTCGCCTCACACCTTCAAATGCATGGAAGGACT	1924
Qy	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer	40
		::::	
Db	1925	CACACTGGAGAAACCCTATGAATGTAAAGCAGTGTCGGAAGCTCTTTTGGATGTGCCTCG	1984
Qy	41	AsnLeuThrLysHisLysLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60
		::::	
Db	1985	CGACTTCAAAATGCATGGAAGGACTTCACACTGGAGAGAAACCGTATAAATGTAAAGCAATGT	2044

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
 Db 2045 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAGGCATGGAAGGACTCAC 2095

RESULT 10						
AC008770/c						
LOCUS	AC008770	178184 bp	DNA	linear	PRI 22-MAR-2003	
DEFINITION	Homo sapiens chromosome 19 clone CTD-2006C1, complete sequence.					
ACCSSION	AC008770					
VERSION	AC008770.7	GI:15887253				
KEYWORDS	HTG.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 178184)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
2 (bases 1 to 178184)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 178184)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (31-AUG-2000) DOE Joint Genome Institute. 2800 Mitchell

REFERENCE	Drive, Walnut Creek, CA 94598, USA
AUTHORS	4 (bases 1 to 178184)
TITLE	DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL	Direct Submission Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	5 (bases 1 to 178184)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Oct 3, 2001 this sequence version replaced gi:9954580. Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.7% of Sequence;

```

FEATURES
  source
    1..178184
      Location/Qualifiers
        STS_Content:
          SHGC-37302 G30664
          SHGC-36253 G28435
          SHGC-14333 G13850
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="19"
        /clone="CTD-2006C1"

```

ORIGIN

Alignment Scores:

Pred. No.:	3,678-25	Length:	178184
Score:	327.00	Matches:	56
Percent Similarity:	81.82%	Conservative:	7
Best Local Similarity:	72.73%	Mismatches:	14
Query Match:	73.98%	Indels:	0
DB:	8	Gaps:	0

US-10-669-861-2 COPY 33 109 (1-77) X AC008770 (1-178184)

Qy	1	CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeulleArgHisGlnArgThr	20
Db	158632	TGTAAGCAGTGTGGGAAAGCCTTTGGATATGCCTTCACACCTTCAAAATGCATGGAGGACT	158573
Qy	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer	40
Db	158572	CACACTGGAGAGAAACCTTATGAATGTAAAGCAGTGTGGGAAGTCTTTTGGATGTGCCTCG	158513
Qy	41	AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60
Db	158512	CGACTTCAATGCATGGAGGACTCACACTGGAGAGAAACCGTATTAATGTATGAAGCAATGT	158453
Qy	61	GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis	77
Db	158452	GGGAAAGCTTTTGGATGTCCCTCAAACTTCGAAGGCATGGAAGACTCAC	158402

[illegible]

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible in this format but are available as part of this entry's ASN.1 file.

----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC

Consensus quality: 179128 bases at least Q20
 Insert size: 210180; 20.5% error; agarose-fp
 Insert size: 179518; sum-of-contigs
 Quality coverage: 7.36x in Q20 bases; agarose-fp
 Quality coverage: 8.62x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 34830: contig of 34830 bp in length
 * 34831 34930: gap of unknown length
 * 34931 94552: contig of 59622 bp in length
 * 94553 94652: gap of unknown length
 * 94653 179581: contig of 84929 bp in length.

FEATURES

source
 1. .179581
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /clone="RP11-197P13"
 /sex="male"
 /tissue type="blood"
 /clone_lib="RP11 human BAC library 11"
 1. .34830
 /notes="assembly_name:Contig61
 clone_end:r7
 vector_side:right"
 34831_-34930
 /estimated_length=unknown
 34931_.94552
 /notes="assembly_name:Contig62"
 94553_.94652
 /estimated_length=unknown
 94653_.179581
 /notes="assembly_name:Contig63
 clone_end:SP6
 vector_side:left"

ORIGIN

Alignment Scores:
 Pred. No.: 3,71e-25 Length: 179581
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC009397 (1-179581)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
 Db 115608 TGTAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAATGCTGAAGGACT 115549
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 115548 CACACTGGAGAGAAACCCCTATGAATGAAGCAGTGTGGGAGTCTTTTGGATGCTCTG 115489
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlnCys 60
 Db 115488 CGACTTCAATGCTGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 115429
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 115428 GGGAAAGCTTTTGGATGCTCCCTCAACCTTGAAGGACTGAAGGACTCAC 115378

RESULT 12

AC022415/c

LOCUS AC022415 180510 bp DNA linear PRI 22-DEC-2000
 DEFINITION Homo sapiens chromosome 19 clone CTC-359D24, complete sequence.
 AC022415
 VERSION AC022415.5 GI:11968305
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 180510)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 180510)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-FEB-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 180510)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (22-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 COMMENT On Dec 22, 2000 this sequence version replaced gi:7704995.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.9% of Sequence;
 Estimated Total Number of Errors is 0.1.
 STS Content:
 SHGC-6150 GI4148.
 Location/Qualifiers
 1..180510
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTC-359D24"

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 3,73e-25 Length: 180510
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: 8 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC022415 (1-180510)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
 Db 8022 TGTAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAATGCTGAAGGACT 7963
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 7962 CACACTGGAGAGAAACCCCTATGAATGAAGCAGTGTGGGAGTCTTTTGGATGCTCTG 7903
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlnCys 60
 Db 7902 CGACTTCAATGCTGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 7843
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 7842 GGGAAAGCTTTTGGATGCTCCCTCAACCTTGAAGGACTGAAGGACTCAC 7792

RESULT 13

AX202430

LOCUS

DEFINITION

AX202430

Sequence 18 from Patent WO0152620.

3300 bp

DNA

linear

PAT 30-AUG-2001

```

VERSION      AX202430.1  GI:15392178
KEYWORDS     .
SOURCE       synthetic construct
ORGANISM     synthetic construct
other sequences; artificial sequences.
REFERENCE    1
AUTHORS      Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE        Methods and compositions to modulate expression in plants
JOURNAL      Patent: WO 0152620-A 18 26-JUL-2001;
              The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
              DISCOVERY, INC. (CA)
FEATURES     Location/Qualifiers
             1..3300
             /organism="synthetic construct"
             /mol_type="unassigned DNA"
             /db_xref="taxon:34630"
             /note="Partial sequence of pMal-Ap3 and zinc finger protein
             ZFPp3"
ORIGIN
Alignment Scores:
Pred. No.:      7,178-27      Length:      3300
Score:          325.00      Matches:      54
Percent Similarity: 84.4%      Conservative: 11
Best Local Similarity: 70.13%      Mismatches:  12
Query Match:    73.53%      Indels:       0
DB:             6           Gaps:         0
US-10-669-861-2_COPY_33_109 (1-77) x AX202430 (1-3300)
QY    1  CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
      |||  :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db    2758  TGTCGGGAATGTTGGTAAGTCCTTCAGCAGCAGCTCCCTGGTGGCCGACCAGGGTACC 2817
QY    21  HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
      |||  :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db    2818  CACACGGGTGAAAAACCGGTATAAATGCCCAGAGTCGGCGCAAAATCTTTAGCCAGTCCAGC 2877
QY    41  AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
      |||  :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db    2878  AACCTGGTGGCCGCACTCAACGCACTCATCTGCGAGAGCCATACAAATGTCAGAAATGT 2937
QY    61  GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
      |||  :::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db    2938  GGCAGTCTTTCAGCCAGCTCCAGCACTCTGGTGGCCACCAACGACTACTC 2988
RESULT 14
AC103247/c
LOCUS
DEFINITION  Rattus norvegicus clone CH230-148G24, WORKING DRAFT SEQUENCE, 3
            unordered pieces.
ACCESSION  AC103247
VERSION    AC103247.5  GI:30578854
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
            1 (bases 1 to 216979)
            Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alibrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Divyak.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

```

```

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guarva,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Longacre,S., Lopez,J.,
Liu,J., Liu,W., Liu,Y., London,P., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Muidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarunpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shateman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,B., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 216979)
Worley,K.C.
Direct Submission
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 216979)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23120050.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJPH
Center clone name: CH230-148G24
----- Summary Statistics
Assembly program: Atlas 3.0;

```

Consensus quality: 211343 bases at least Q40
Consensus quality: 212467 bases at least Q30
Consensus quality: 213063 bases at least Q20
Estimated insert size: 219835; sum-of-coverage estimation
Quality coverage: 10x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 3 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

* 1 55572: contig of 55572 bp in length
* 55573 55672: gap of unknown length
* 55673 215556: contig of 159884 bp in length
* 215557 215656: gap of unknown length
* 215657 216979: contig of 1323 bp in length.

FEATURES
source
1. 216979
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-148G24"
50131..50826
/notes="clone boundary
clone end:T7
site:ECORI
55573..55672
/estimated_length=unknown
215557..215656
/estimated_length=unknown

misc_feature
1.7e-24 Length: 216979
322.00 Matches: 55
85.71% Conservatives: 11
71.43% Mismatches: 11
72.85% Indels: 0
14 Gaps: 0

gap
end_sequence:RWBAN48TJB"
55573..55672
/estimated_length=unknown
215557..215656
/estimated_length=unknown

gap
end_sequence:RWBAN48TJB"
55573..55672
/estimated_length=unknown
215557..215656
/estimated_length=unknown

ORIGIN
Alignment Scores:
Pred. No.: 1.7e-24 Length: 216979
Score: 322.00 Matches: 55
Percent Similarity: 85.71% Conservatives: 11
Best Local Similarity: 71.43% Mismatches: 11
Query Match: 72.85% Indels: 0
DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC103247 (1-216979)
Qy 1 CyslyAspCysGlylyAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 2995 TGGCGTGACTGTGCCAAAGCCTTCGGCGACGCTCAGCCCTTATTCGGCACCAGAGACA 2936
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 2935 CACACTGGAGAGAGCCTTATAAATCAGAGACTGTGGAAGAGCCTTCAGCCAGAGCTCA 2876
Qy 41 AsnLeuThrlyHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 2875 TCTCTTACAAAACACAGAAAACCTCACACTGGGAAGACCTTAATAGTGTAAGAATGT 2816
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuAArgHisGlyArgThrHis 77
Db 2815 GGAAAAGCGTTTAGCCAGAGTTCATCTCTTCTCAACATCAGAAAATCAT 2765

RESULT 15
AC133610/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-70117, *** SEQUENCING IN PROGRESS
AC133610
AC133610.2 GI:25073277
HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
ACCESSION
VERSION
KEYWORDS
SOURCE

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genomex Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KGIN
Center clone name: CH230-70117
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 201502 bases at least Q40
Consensus quality: 204933 bases at least Q30
Consensus quality: 207525 bases at least Q20
Estimated insert size: 207494; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 109488: contig of 109488 bp in length
* 109489 109588: gap of unknown length
* 109589 237009: contig of 127421 bp in length
* 237010 237109: gap of unknown length
* 237110 238307: contig of 1198 bp in length.

FEATURES
source
1..238307
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-70117"
40706..41956
/note="wgs_contig"
106919..109488
/note="wgs_contig"
109489..109588
/estimated_length=unknown
109589..110614
/note="wgs_contig"
237010..237109
/estimated_length=unknown

ORIGIN
Alignment Scores:
Pred. No.: 1.89e-24 Length: 238307
Score: 322.00 Matches: 55
Percent Similarity: 85.71% Conservative: 11
Best Local Similarity: 71.43% Mismatches: 11
Query Match: 72.85% Indels: 0
DB: 14 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AC133610 (1-238307)
Qy 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 2995 TCGGTGACTGTGCCAAGCCCTTCGGCACCGCTCAGCCCTTATTCCGACAGAGAAC 2936
Qy 21 HisThrGlyGluLysProTyrLysCysGluCysGlyLysAlaPheThrGlnSerSer 40

Db 2935 CACACTGGAGAGAGCCTTATAAATGCAAGGACTGTGGAAGGCTTTTCAGCCAGAGCTCA 2876
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysGlyGlnCys 60
Db 2875 TCTCTTACAAACACACAGAAACTTCACACTGGAGAGAGCCCTATAAGTGTAAAGAATGT 2816
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
Db 2815 GGAAGAGCGTTTACCCAGAGTTTCATCTCTTCTCAACATCAGAAAAATCAT 2765

Search completed: January 12, 2006, 02:31:37
Job time : 3742 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 00:35:34 ; Search time 462 Seconds
(without alignments)
1110.782 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCKAFIQKSLNRHQRT.....KCGKAFGCPNLRHRGRTH 77

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cpn2_1/USPTO/spool/US10669861/runat_09012006_144452_20361/app_query.fasta_1.263
-DB=N_Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10669861@CCN 1 1 727 @runat_09012006_144452_20361 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 21.*

1:	Geneseq1980s.*
2:	Geneseq1990s.*
3:	Geneseq2000s.*
4:	Geneseq2001as.*
5:	Geneseq2001bs.*
6:	Geneseq2002as.*
7:	Geneseq2002bs.*
8:	Geneseq2003as.*
9:	Geneseq2003bs.*
10:	Geneseq2003cs.*
11:	Geneseq2003ds.*
12:	Geneseq2004as.*
13:	Geneseq2004bs.*
14:	Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	1137	11 ADL33661	Adl33661 Zinc fing
2	329	74.4	3569	4 ABA07313	ABA07313 Human pan
3	329	74.4	3569	4 AAK90472	Aak90472 Human dig
4	327	74.0	672	6 ABR71555	Abk71555 Human dit

ALIGNMENTS

RESULT 1

ADL33661

ID ADL33661 standard; DNA; 1137 BP.

AC ADL33661;

DT 20-MAY-2004 (first entry)

DE Zinc finger domain-containing protein (Neurol-p65) coding sequence.

KW modified cell; artificial transcription factor; stress resistance;

KW phenotypic trait alteration; zinc finger domain; gene; ds.

OS Unidentified.

XX WO2003048345-A1.

XX PD 12-JUN-2003.

XX PF 07-DEC-2002; 2002WO-KR002309.

XX PR 07-DEC-2001; 2001US-0338441P.

XX PR 26-APR-2002; 2002US-0376053P.

XX PR 02-AUG-2002; 2002US-0400904P.

XX PR 05-AUG-2002; 2002US-0401089P.

XX (TOOL-) TOOLGEN INC.

Aas86929 DNA encod
Adn99083 Novel hum
Ado0652 Novel hum
Aas18786 cDNA enco
Aas18792 Human mdd
Abx34635 Human mdd
Aad41202 Human nuc
Ado03740 Human cDN
Ado00362 Novel hum
Adn98793 Novel hum
Adn99082 Novel hum
Ado00651 Novel hum
Adg83562 Human tum
Adl33677 Zinc fing
Aad11592 Partial s
Aad11613 Zinc fing
Aad11588 Partial s
Aad11589 Partial s
Aas3064 Human pol
Aas82573 DNA encod
Aas2080 Human pol
Adm02541 Human cDN
Ach68852 Human gen
Aca98974 cDNA enco
Ach15895 Human adu
Adl33669 Zinc fing
Aah05264 Human cDN
Ach92086 Human gen
Aah15787 Human cDN
Aai85395 Human pol
Adc30336 Human nov
Adb63310 Human cDN
Aca98925 cDNA enco
Adm01826 Human cDN
Adg25219 Human sof
Acc46347 Human dit
Aas64586 DNA encod
Adt07945 Full leng
Abk83826 Human cDN
Adm04207 Antipsori
Adr25049 Breast ca

5 AAS86929 1668 5
6 1746 12
7 1746 12
8 1746 12
9 1765 6
10 2192 6
11 2217 8
12 2275 6
13 2280 11
14 2718 12
15 2718 12
16 2718 12
17 3825 13
18 636 11
19 3300 4
20 330 4
21 3300 4
22 3300 4
23 2056 4
24 2175 5
25 2496 4
26 2667 11
27 505 12
28 2115 9
29 468 9
30 630 11
31 827 4
32 1258 12
33 1449 4
34 1512 4
35 1890 10
36 1914 10
37 2257 9
38 2664 11
39 2890 12
40 3309 8
41 3639 5
42 3832 13
43 3839 6
44 3839 12
45 3839 13

c

PI Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;
PI Jang Y;
XX WPI; 2003-513760/48.
DR P-PSDB; ADL33662.
XX
XX New modified cell comprising a heterologous nucleic acid encoding an
PT artificial transcription factor that confers stress resistance, useful
PT for altering a phenotypic trait of a cell or organism.
XX
XX Disclosure; SEQ ID NO 201; 169pp; English.
XX
CC The invention comprises a modified cell containing a heterologous nucleic
CC acid encoding an artificial transcription factor that confers stress
CC resistance to the modified cell. The modified cell of the invention is
CC useful for altering a phenotypic trait of a cell or organism. The present
CC DNA sequence encodes a protein which contains zinc finger domains.
XX
SQ Sequence 1137 BP; 280 A; 373 C; 269 G; 215 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.99e-45 Length: 1137
Score: 442.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps:

US-10-669-861-2_COPY_33_109 (1-77) x ADL33661 (1-1137)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAtgHicGlnArgThr 20
DB 97 TGTAAAGATTGGGAAAGCTTTTCATTGAGAGTCAAAACCTCATGACACACGAGAACT 156

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 157 CACACCGGGGAAAAACCGTACAGGTGTGAAGATGTGGCAAAGCTTTTACCCAACTCTCA 216

QY 41 AsnLeuThrLysHisLysLysLysIleHisThrGlyGluLysProTyrLysCysGlnCys 60
DB 217 AACCTTACTAACAATAAGAAAAATTCATACCGGGGAAAAACCGTATAATGTGAAGCAATGT 276

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 277 GGGAAAGCTTTTGGATGTCCCTCAACCTTCGAAGGATGAAGACTCAC 327

RESULT 2
ID ABA07313
XX ABA07313 standard; DNA; 3569 BP.
AC ABA07313;
XX
DT 14-JAN-2002 (first entry)
XX
DE Human pancreatic cancer related genomic DNA, SEQ ID NO: 632.
XX
KW Human; cytostatic; antidiabetic; antiinflammatory; gastric; osteopathic;
KW antihormone; antiulcer; thyroid-active; gene therapy; antisense therapy;
KW pancreatic cancer antigen inhibitor; pancreatic cancer; pancreatitis;
KW diabetes; endocrine disorder; acromegaly; hyperthyroidism;
KW gastrointestinal disorder; Crohn's disease; duodenal ulcer; ds.
XX
OS Homo sapiens.
XX
PN WO200155206-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US0001353.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209457P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218292P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.

02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254907P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-457717/49.
 XX
 PT Isolated pancreatic cancer polypeptide for treating, preventing and/or
 PT prognosing disorders related to the pancreas including pancreatic cancers
 PT and also for testing and detection e.g. diagnosis.
 XX
 PS Disclosure; SEQ ID NO 632; 537pp; English.
 XX

CC The invention relates to an isolated polypeptide comprising an amino acid
 CC sequence at least 90% identical to 188 amino acid sequences fully defined
 CC in the specification and encoded by 188 cDNA clones fully defined in the
 CC specification. The invention also relates to a fragment having biological
 CC activity, a domain, an epitope, full length protein, variant, allelic
 CC variant or a species homologue of the fully defined sequence. The
 CC polynucleotide and polypeptide are useful for treating, preventing and/or
 CC prognosing disorders related to the pancreas including pancreatic cancer,
 CC pancreatitis, diabetes, endocrine disorders such as acromegaly or
 CC hyperthyroidism, and gastrointestinal disorders such as Crohn's disease
 CC and duodenal ulcers. The present sequence encodes a pancreatic cancer-
 CC related polypeptide of the invention
 XX
 SQ Sequence 3569 BP; 1193 A; 673 C; 773 G; 929 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 1.32e-30 Length: 3569
 Score: 329.00 Matches: 54
 Percent Similarity: 84.42% Conservative: 11
 Best Local Similarity: 70.13% Mismatches: 12
 Query Match: 74.43% Indels: 0
 DB: 4 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x ABA07313 (1-3569)
 QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 DB 2436 TCGAAGCAATGTGGTAAGCCCTTCATTCCATTCAGTCCCTCGATATCATGAAGACT 2495
 QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
 DB 2496 CACACTGGAGAGAAACCCCTATGAGTGAAGCAATGTGGAGAGCCCTTCAGATCTGCCTCA 2555
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 DB 2556 CACCTTCGAATCGTAAAGGACTCACACTGGAGAGAAACCCCTATGAGTGAAGCAATGT 2615
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 DB 2616 GGGAAAGCCTTCAGTTGTGCTCAACCTTCGAAGAGCATGGTAGACTCAC 2666
 RESULT 3
 AAK90472
 ID AAK90472 standard; DNA; 3569 BP.
 XX
 AC AAK90472;
 XX
 DT 05-NOV-2001 (first entry)
 XX Human digestive system antigen genomic sequence SEQ ID NO: 4048.
 DE Human; digestive system antigen; gene therapy; cancer; appendicitis;
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 KW digestive system disorder; Meckel's diverticulum; ds.
 XX Homo sapiens.
 XX
 XX WO20015314-A2.
 PN
 XX
 PD 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US001324.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226273P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0233080P.
PR 08-SEP-2000; 2000US-0233081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 4048; 986bp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention

```

XX
SQ Sequence 3569 BP; 1193 A; 673 C; 773 G; 929 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 1,32e-30 Length: 3569
Score: 329.00 Matches: 54
Percent Similarity: 84.42% Conservatve: 11
Best Local Similarity: 70.13% Mismatches: 12
Query Match: 74.43% Indels: 0
DB: 4 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AAK90472 (1-3569)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
Db 2436 TGCAGCAATGTGTAAAGCCTTCATTTCAGTTCCTTCGATATCATGAAAGGACT 2495

Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 2496 CACACTGGAGAAACCCCTATGAGTGTAAAGCAATGTGGGAAGGCCCTTCAGATCTGCCTCA 2555

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 2556 CACCTTCGAATGCATGAAGGACTCACACTGGAGAGAAACCTATGAGTGTAAAGCAATGT 2615

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuAArgHisGlyArgThrHis 77
Db 2616 GGGAAAGCCTTCAGTTGTGCCTCAACCTTCGNAAGCATGTAGGACTCAC 2666

RESULT 4
ABK71555
ID ABK71555 standard; cDNA; 672 BP.
XX
AC ABK71555;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human dithp polynucleotide #21.
XX
KW Human; dithp; diagnostic and therapeutic polynucleotide; gene; ss; bone;
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
KW inflammatory disorder; viral infection; bacterial infection; seizure;
KW fungal infection; parasitic infections; developmental disorder; breast;
KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
KW gastrointestinal disorder; transport disorder; gene therapy; kidney;
KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
KW skin; testis; thymus.
XX
OS Homo sapiens.
XX
PN WO200220754-A2.
XX
PD 14-MAR-2002.
XX
PF 29-AUG-2001; 2001WO-US027127.
XX
PR 05-SEP-2000; 2000US-0229747P.
PR 05-SEP-2000; 2000US-0229748P.
PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.
PR 05-SEP-2000; 2000US-0230583P.
PR 05-SEP-2000; 2000US-0230505P.
PR 06-SEP-2000; 2000US-0230514P.
PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.
PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.
PR 06-SEP-2000; 2000US-0230595P.
PR 06-SEP-2000; 2000US-0230597P.
PR 06-SEP-2000; 2000US-0230598P.
PR 06-SEP-2000; 2000US-0230599P.
PR 06-SEP-2000; 2000US-0230610P.

```

PR	06-SEP-2000; 2000US-0230865P.	
PR	06-SEP-2000; 2000US-0230988P.	
PR	07-SEP-2000; 2000US-0230951P.	
PR	07-SEP-2000; 2000US-0231163P.	
PR	07-SEP-2000; 2000US-0231167P.	
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
XX	Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;	
PI	Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;	
PI	Montiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;	
PI	Gerstlin ER, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;	
PI	Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR;	
XX		
DR	WPI: 2002-383054/41.	
DR	P-FSDB; ABG59963.	
XX		
XX	An isolated polynucleotide useful in diagnostics and therapeutics.	
PT	Claim 1; Page 415; 686pp; English.	
PS		
XX		
XX	The invention relates to human diagnostic and therapeutic (dithp)	
CC	polynucleotides and their associated polypeptides (DITHP polypeptides).	
CC	The sequences of the invention are used in the treatment and diagnosis of	
CC	cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers	
CC	(e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,	
CC	cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or	
CC	thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,	
CC	psoriasis, osteoporosis), viral infections, bacterial infections, fungal	
CC	infections, parasitic infections, developmental disorders (e.g. anaemia,	
CC	epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),	
CC	endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders	
CC	(e.g. obesity, diabetes), neurological disorders (e.g. stroke,	
CC	amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal	
CC	disorders (e.g. ulcerative colitis, lysinuria) and transport disorders	
CC	(e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences	
CC	ABK71535-ABK71809 represent human dithp polynucleotides of the invention	
XX		
SQ	Sequence 672 BP; 213 A; 137 C; 163 G; 159 T; 0 U; 0 Other;	
Align. No.:	2,568-31	Length: 672
Score:	327.00	Matches: 56
Percent Similarity:	81.82%	Conservative: 7
Best Local Similarity:	72.73%	Mismatches: 14
Query Match:	73.98%	Indels: 0
DB:	6	Gaps: 0

US-10-669-861-2 COPY 33 109 (1-77) x ABK71555 (1-672)

Qy	1	CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr	20
Db	311	TGTAAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGATCGAAGGACT	370
Qy	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer	40
Db	371	CACACTGGAGAGAAACCTTATGAATGAACGACGTGTGGAGTCTTTTGGATGTGCCTCG	430
Qy	41	AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60
Db	431	CGACTTCCTCAATGCATGGNAGGACTCACACTGGAGAGAAACCGTATAAATGTAAAGCAATGT	490
Qy	61	GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis	77
Db	491	GGGAAAGCTTTTGGATGTCCCTCAAACTTCGAAGGCATGGAAGGACTCAC	541

RESULT 5	
AAS86929	
ID	AAS86929 standard; cDNA; 1668 BP.
XX	
AC	AAS86929;
XX	
DT	13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22733.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX
XX WPI; 2001-639362/73.
DR
XX P-PSDB; ABG22742.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
PT
XX
XX Claim 1; SEQ ID NO 22733; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (I). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1668 BP; 555 A; 334 C; 389 G; 390 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,588-31 Length: 1668
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservativity: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 5 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AAS86929 (1-1668)

QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
DB 595 TGTAAAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 654

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrClnSerSer 40
DB 655 CACACTGGAGAGAAACCCATATGAATGAACGAGTGTGGGAAGTCTTTGGATGTGCCCTCG 714

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
DB 715 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAAACCGTATATAATGTAAGCAATGT 774

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
DB 775 GGGAAAGCTTTGGATGTCCCTCAAAAGCTTCGAAGGCATGGAAGGACTCAC 825

RESULT 6
ADN99083
ID ADN99083 standard; cDNA; 1746 BP.
XX AC ADN99083;
XX
XX 29-JUL-2004 (first entry)
XX
XX Novel human cDNA sequence #683.
KW ds; gene; anti-inflammatory; dermatological; neuroprotective;
KW immunomodulator; antibacterial; virucide; antipsoriatic; cytostatic;
KW gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
KW psoriasis; diabetes; early aging; hormonal imbalance;
KW ischemic heart disease; ulcerative colitis.
XX
XX Homo sapiens.
XX WO2004038003-A2.
XX
XX 06-MAY-2004.
XX
XX 24-OCT-2003; 2003WO-US033947.
XX
XX 25-OCT-2002; 2002US-0421061P.
XX 25-OCT-2002; 2002US-0421080P.
XX 25-OCT-2002; 2002US-0421552P.
XX 25-OCT-2002; 2002US-0421614P.
XX 30-OCT-2002; 2002US-0422177P.
XX 30-OCT-2002; 2002US-0422178P.
XX 15-NOV-2002; 2002US-0426355P.
XX 15-NOV-2002; 2002US-0426384P.
XX 15-NOV-2002; 2002US-0426394P.
XX 15-NOV-2002; 2002US-0426430P.
XX 15-NOV-2002; 2002US-0426916P.
XX 27-NOV-2002; 2002US-0429224P.
XX 27-NOV-2002; 2002US-0429275P.
XX 27-NOV-2002; 2002US-0429302P.
XX 27-NOV-2002; 2002US-0429326P.
XX 04-DEC-2002; 2002US-0429651P.
XX 04-DEC-2002; 2002US-0430645P.
XX 04-DEC-2002; 2002US-0430651P.
XX 04-DEC-2002; 2002US-0430657P.
XX 04-DEC-2002; 2002US-0430663P.
XX 04-DEC-2002; 2002US-0430668P.
XX 05-DEC-2002; 2002US-0430937P.
XX 05-DEC-2002; 2002US-0430965P.
XX 05-DEC-2002; 2002US-0431458P.
XX 12-DEC-2002; 2002US-0433251P.
XX 12-DEC-2002; 2002US-0433500P.
XX 13-DEC-2002; 2002US-0433316P.
XX 13-DEC-2002; 2002US-0433318P.
XX 23-DEC-2002; 2002US-0436236P.
XX 03-JAN-2003; 2003US-0437914P.
XX 17-JAN-2003; 2003US-0440820P.
XX 17-JAN-2003; 2003US-0440821P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.

19-MAY-2003; 2003US-04711306P.
 19-MAY-2003; 2003US-04711336P.
 22-MAY-2003; 2003US-0472420P.
 22-MAY-2003; 2003US-0472430P.
 09-JUN-2003; 2003US-0476609P.
 09-JUN-2003; 2003US-0476621P.
 09-JUN-2003; 2003US-0476632P.
 09-JUN-2003; 2003US-0476641P.
 08-JUL-2003; 2003US-0485217P.
 08-JUL-2003; 2003US-0485218P.
 08-JUL-2003; 2003US-0485223P.
 08-JUL-2003; 2003US-0485224P.
 08-JUL-2003; 2003US-0485325P.
 08-JUL-2003; 2003US-0485359P.
 14-JUL-2003; 2003US-0486446P.
 14-JUL-2003; 2003US-0486480P.
 15-JUL-2003; 2003US-0486891P.
 15-JUL-2003; 2003US-0486960P.
 08-AUG-2003; 2003US-0493341P.
 08-AUG-2003; 2003US-0493370P.
 08-AUG-2003; 2003US-0493573P.
 08-AUG-2003; 2003US-0493577P.
 (FIVE-) FIVE PRIME THERAPEUTICS INC.
 Williams LT, Chu K, Lee E, Heesters P, Behrens D;
 Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
 Wong JGP, Wu G, Zhang H, Zeng C;
 WPI; 2004-365511/34.
 P-PSDB; ADM99867.
 New nucleic acid molecules, useful in preparing a composition for
 treating or preventing e.g. inflammatory CNS, bacterial or viral
 disorders, cancer, psoriasis, diabetes, ischemic heart disease or
 ulcerative colitis.
 Claim 1; SEQ ID NO 683; 532pp; English.
 The invention relates to a nucleic acid molecule comprising a
 polynucleotide sequence or its complement that encodes a polypeptide. The
 nucleic acid is useful in preparing a composition for treating or
 preventing inflammatory, CNS, immune, bacterial or viral disorder,
 cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
 heart disease or ulcerative colitis. This sequence corresponds to a
 nucleic acid of the invention.
 Sequence 1746 BP; 595 A; 349 C; 374 G; 428 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9, 12e-31 Length: 1746
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: 12 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x ADM99083 (1-1746)
 QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
 1414 TGTAAAGCAGTGTGGGAAGCCCTTGGATCGCTCACACCTTCAAATGCATGGAAGGACT 1473
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 1474 CACACTGGAGAGAACCCCTATGAAGTAAAGCTGGGAAGTCTTTGGATGTGCCTCG 1533
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 1534 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAACCCCTATGAAGTAAATGTGAAGCAATGT 1593
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 |||||

Db 1594 GGGAAAGCTTTTGGATGTCCTCAACCTTCGAAGCAGTGAAGGACTCAC 1644
 RESULT 7
 ADO00652
 ID ADO00652 standard; cDNA; 1746 BP.
 XX
 AC ADO00652;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Novel human cDNA sequence #1467.
 XX
 KW ds; gene; anti-inflammatory; dermatological; neuroprotective;
 immunomodulator; antibacterial; virucide; antiproliferative; cytostatic;
 gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
 psoriasis; diabetes; early aging; hormonal imbalance;
 ischemic heart disease; ulcerative colitis.
 XX
 OS Homo sapiens.
 XX
 PN WO2004038003-A2.
 XX
 PD 06-MAY-2004.
 XX
 PF 24-OCT-2003; 2003WO-US033947.
 XX
 PR 25-OCT-2002; 2002US-0421061P.
 PR 25-OCT-2002; 2002US-0421080P.
 PR 25-OCT-2002; 2002US-0421552P.
 PR 25-OCT-2002; 2002US-0421614P.
 PR 30-OCT-2002; 2002US-0422177P.
 PR 30-OCT-2002; 2002US-0422178P.
 PR 15-NOV-2002; 2002US-0426355P.
 PR 15-NOV-2002; 2002US-0426384P.
 PR 15-NOV-2002; 2002US-0426394P.
 PR 15-NOV-2002; 2002US-0426430P.
 PR 15-NOV-2002; 2002US-0426916P.
 PR 27-NOV-2002; 2002US-0429224P.
 PR 27-NOV-2002; 2002US-0429275P.
 PR 27-NOV-2002; 2002US-0429302P.
 PR 27-NOV-2002; 2002US-0429326P.
 PR 27-NOV-2002; 2002US-0429651P.
 PR 04-DEC-2002; 2002US-0430645P.
 PR 04-DEC-2002; 2002US-0430651P.
 PR 04-DEC-2002; 2002US-0430657P.
 PR 04-DEC-2002; 2002US-0430663P.
 PR 04-DEC-2002; 2002US-0430668P.
 PR 04-DEC-2002; 2002US-0430684P.
 PR 05-DEC-2002; 2002US-0430937P.
 PR 05-DEC-2002; 2002US-0430965P.
 PR 05-DEC-2002; 2002US-0430965P.
 PR 12-DEC-2002; 2002US-0433251P.
 PR 12-DEC-2002; 2002US-0433500P.
 PR 13-DEC-2002; 2002US-0433316P.
 PR 13-DEC-2002; 2002US-0433319P.
 PR 23-DEC-2002; 2002US-0436236P.
 PR 03-JAN-2003; 2003US-0437914P.
 PR 17-JAN-2003; 2003US-0440820P.
 PR 17-JAN-2003; 2003US-0440821P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 18-APR-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 02-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 19-MAY-2003; 2003US-0471336P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 22-MAY-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 09-JUN-2003; 2003US-0476621P.

```
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RP, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI; 2004-365511/34.
XX P-PSDB; ADN99867.
XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
XX Claim 1; SEQ ID NO 2251; 532pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.
XX
XX Sequence 1746 BP; 595 A; 349 C; 374 G; 428 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 9,128-31 Length: 1746
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 12 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x ADO00652 (1-1746)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 1414 TGTAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 1473
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1474 CACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAAGTCTTTGGATGTGCCTCG 1533
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
DB 1534 CGACTTCAATGTCGGAAGGACTCACACTGGAGAGAACCGTATAAATGTAAGCAATGT 1593
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 1594 GGGAAAGCTTTTGGATGTCCCTCAACCTTCGAAGGCATGGNAGGACTCAC 1644
RESULT 8
AAS18786
ID AAS18786 standard; cDNA; 1765 BP.
XX
```

```
AC AAS18786;
XX
XX 26-MAR-2002 (first entry)
XX
XX cDNA encoding human zinc finger protein 53.
XX
XX Human; zinc finger protein 53; malignant tumour; HIV infection;
XX immunological disease; inflammation; 88.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 189..1640
XX FT /*tag= a
XX FT /product= "Zinc finger protein 53"
XX
XX CNI307009-A.
XX
XX 08-AUG-2001.
XX
XX 21-JAN-2000; 2000CN-00111472.
XX
XX 21-JAN-2000; 2000CN-00111472.
XX (BODA-) BODAO GENE TECH CO LTD SHANGHAI.
XX
XX Mao Y, Xie Y;
XX WPI; 2002-049871/07.
XX DR P-PSDB; AAU10794.
XX
XX Polypeptide-human zinc finger protein 53 and polynucleotide for coding
XX said polypeptide.
XX
XX Claim 6; Page 25-26 (Disclosure); 37pp; Chinese.
XX
XX The present invention relates to the isolation of human zinc finger
XX protein 53 and the cDNA sequence encoding it. Also described is the DNA
XX recombination process to produce zinc finger protein 53, and methods of
XX applying zinc finger protein 53 and the polynucleotide encoding it in
XX treating various diseases, such as malignant tumours, HIV infection,
XX immunological diseases and inflammations. The present invention also
XX discloses the antagonist resisting the polypeptide and its treatment
XX effect. The present sequence encodes for the human zinc finger protein 53
XX of the invention
XX
XX Sequence 1765 BP; 597 A; 354 C; 383 G; 431 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 9,256-31 Length: 1765
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0
US-10-669-861-2_COPY_33_109 (1-77) x AAS18786 (1-1765)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 1308 TGTAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 1367
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 1368 CACACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAAGTCTTTGGATGTGCCTCG 1427
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
DB 1428 CGACTTCAATGTCGGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 1487
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 1488 GGGAAAGCTTTTGGATGTCCCTCAACCTTCGAAGGCATGGAAGGACTCAC 1538
```


RESULT 9
 ABS51792
 ID ABS51792 standard; cDNA; 2192 BP.
 XX
 AC ABS51792;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human mddt cDNA Incyte ID No: LI:1175131.1:2001JAN12.
 XX
 KW Human; molecule for disease detection and treatment; MDDT; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
 KW hepatotropic; immunosuppressive; antiasthmatic; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200255738-A2.
 XX
 XX 18-JUL-2002.
 XX
 XX 09-JAN-2002; 2002WO-US001008.
 XX
 PR 12-JAN-2001; 2001US-0261622P.
 PR 16-JAN-2001; 2001US-0261865P.
 PR 17-JAN-2001; 2001US-0262208P.
 PR 17-JAN-2001; 2001US-0262209P.
 PR 17-JAN-2001; 2001US-0262326P.
 PR 19-JAN-2001; 2001US-0263063P.
 PR 19-JAN-2001; 2001US-0263065P.
 PR 19-JAN-2001; 2001US-0263329P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
 PI Dam TC, Liu TF, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ;
 PI Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;
 XX
 DR WPI; 2002-590679/63.
 DR P-PSDB; ABG70319.
 XX
 PT New disease detection and treatment molecule (MDDT) polynucleotides and
 PT polypeptides, useful in diagnosing, studying, preventing or treating
 PT diseases associated with MDDT expression, e.g. autoimmune or inflammatory
 PT disorders.
 XX
 PS Claim 1; Page 93; 129pp; English.
 XX
 CC The present invention relates to the isolation of novel human molecules
 CC for disease detection and treatment (MDDT), and the polynucleotide
 CC sequences (mddt) encoding them. The MDDT polypeptides may be used to
 CC screen for molecules that bind to, or are bound by the encoded
 CC polypeptides, and to develop a transcript image of a tissue or cell type.
 CC Probes comprising at least 20 nucleotides of the mddt polynucleotide may
 CC be used to assess the toxicity of a test compound. The MDDT polypeptides
 CC and mddt polynucleotides are useful in the diagnosis, study, prevention
 CC and treatment of diseases associated with the expression of molecules for
 CC disease detection and treatment. Such disorders include cell
 CC proliferative disorders (e.g. arteriosclerosis, cirrhosis, or cancers),
 CC and autoimmune/inflammatory disorders (e.g. asthma, Crohn's disease, or
 CC multiple sclerosis). The mddt polynucleotides may also be used as
 CC molecule markers, in microarrays, and in somatic or germline gene
 CC therapy. ABS51779-ABS51814 encode the MDDT proteins of the invention
 XX
 SQ Sequence 2192 BP; 709 A; 450 C; 503 G; 528 T; 0 U; 2 Other;

Query Match: 73.98% Indels: 0
 DB: 6 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x ABS51792 (1-2192)
 QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 DB 1845 TGTAAAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGATGGAAGGACT 1904
 QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 DB 1905 CACACTGGAGAGAAACCCCTATCAATGTAAAGCAGTGTGGGAAGTCTTTTGGATGTCCTCG 1964
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 DB 1965 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAAACCCGTATAAATGTAAGCAATGT 2024
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 DB 2025 GGGAAAGCTTTTGGATGTCCCTCAAACTTCGAAGCATGGAAGGACTCAC 2075
 RESULT 10
 ABX34635
 ID ABX34635 standard; cDNA; 2217 BP.
 XX
 AC ABX34635;
 XX
 DT 13-FEB-2003 (first entry)
 XX
 DE Human mddt cDNA SEQ ID 196.
 XX
 KW MDDT; human; disease detection and treatment molecule polypeptide;
 KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
 KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
 KW gene therapy; protein replacement therapy; cell proliferative disorder;
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
 KW psoriasis; hepatitis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200279449-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-US009944.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-029976P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 DR WPI; 2003-058431/05.
 DR P-PSDB; ABU11645.
 XX
 PT New purified disease detection and treatment molecule proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing cancers
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
 PT or hepatitis.

XX PS Claim 1; SEQ ID NO 196; 339pp + Sequence Listing; English.

XX CC This invention describes a novel disease detection and treatment molecule

XX CC polypeptide (MDDR) which has anti-inflammatory, immunosuppressive,

XX CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,

XX CC antianemic, antipsoriatic and hepatotropic activity. The polynucleotides

XX CC and the polypeptides of the invention can be used for gene therapy,

XX CC protein replacement therapy and are useful for treating a variety of

XX CC diseases or conditions. These polypeptides or polynucleotides are

XX CC particularly useful for diagnosing, treating or preventing cell

XX CC proliferative disorders (e.g. cancers including adenocarcinoma,

XX CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's

XX CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's

XX CC syndrome, inflammation, osteoporosis, thrombocytopaenia, psoriasis or

XX CC hepatitis. ABX34440-ABX344835 encode the MDDR polypeptides represented in

XX CC ABU11450-ABU11845, described in the disclosure of the invention. NOTE:

XX CC The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2217 BP; 711 A; 455 C; 521 G; 530 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.25e-30	Length:	2217
Score:	327.00	Matches:	56
Percent Similarity:	81.82%	Conservative:	7
Best Local Similarity:	72.73%	Mismatches:	14
Query Match:	73.98%	Indels:	0
DB:	8	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x ABX34635 (1-2217)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20

Db 1856 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGATGGAAGGACT 1915

Qy 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40

Db 1916 CACACTGGAGAGAAACCTTATGAATGTAAGCAGTGTGGAAAGCTTTTGGATGTGCCCTCG 1975

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60

Db 1976 CGACTTCAAAATGATGGAAGGACTCACACTGGAGAGAAACCTTATGAATGTAAGCAGTGT 2035

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77

Db 2036 GGGAAAGCTTTTGGATGTCCCTCAAAACCTTGAAGGATGGAAGGACTCAC 2086

RESULT 11

AAD41202

ID AAD41202 standard; cDNA; 2275 BP.

XX AC AAD41202;

XX DT 30-OCT-2002 (first entry)

XX DE Human nucleic acid-associated protein (NAAP-12) cDNA.

XX KW Human; nucleic acid-associated protein; NAAP-12; neurological disorder;

XX KW arteriosclerosis; cancer; cell proliferative disorder; arteriosclerosis;

XX KW lymphoma; epilepsy; Alzheimer's disease; developmental; anticonvulsant;

XX KW autoimmune disorder; AIDS; allergy; anaemia; stroke; malaria; leishmania;

XX KW gene therapy; neurotropic; neuroprotective; cerebroprotective; virucide;

XX KW immunosuppressive; protozoacide; antimicrobial; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 234..2246

XX FT /*tag= a

XX FT /product= "Human NAAP-12"

XX

PN WO200250279-A2.

XX PD 27-JUN-2002.

XX PF 19-DEC-2001; 2001WO-US050256.

XX PR 21-DEC-2000; 2000US-0257714P.

XX PR 05-JAN-2001; 2001US-0260081P.

XX PR 16-JAN-2001; 2001US-0262302P.

XX PR 23-JAN-2001; 2001US-0263823P.

XX PR 02-FEB-2001; 2001US-0266088P.

XX PR 29-OCT-2001; 2001US-0348442P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Baughin MR, Lu Y, Arvizu C, Ramkumar J, Yao MG, Policky JL;

XX PI Wallia NK, Tribouley KW, Yue H, Batra S, Ding L, Lal PG;

XX PI Borowsky ML, Lu DAM, Gandhi AR, Griffin JA, Xu Y, Azimzai Y;

XX PI Gietzen KJ, Tang VT, Warren BA, Mason PM, Burford N, Hafalia AJA;

XX PI Lee EA, Yang J, Gorvad AE, Emerling BM, Marquis JP, Lee SY,

XX PI Swarnakar A, Reddy R;

XX DR WPI; 2002-519887/55.

XX DR P-PSDB; AAE25293.

XX PT Nucleic acid associated proteins and nucleic acids for diagnosing,

XX PT treating and preventing cell proliferative (e.g. cancers), neurological

XX PT (e.g. epilepsy or stroke) or autoimmune disorders (e.g. AIDS).

XX PS Claim 83; Page 188; 193pp; English.

XX CC The invention relates to nucleic acid-associated proteins (NAAP) and

XX CC nucleic acids. The nucleic acid and amino acid sequences are useful for

XX CC diagnosing, treating and preventing cell proliferative e.g.

XX CC arteriosclerosis, atherosclerosis, lymphoma or cancers), neurological

XX CC (e.g. epilepsy, Alzheimer's disease or stroke), developmental, and

XX CC autoimmune disorders (e.g. AIDS, allergies, or anaemia) or infections

XX CC (e.g. malaria, or leishmania), as well as in assessing the effects of

XX CC exogenous compound on the expression of nucleic acid and amino acid

XX CC sequences of nucleic acid-associated proteins. The invention is useful in

XX CC gene therapy. The present sequence is human NAAP-12 cDNA

XX SQ Sequence 2275 BP; 716 A; 473 C; 548 G; 538 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.3e-30	Length:	2275
Score:	327.00	Matches:	56
Percent Similarity:	81.82%	Conservative:	7
Best Local Similarity:	72.73%	Mismatches:	14
Query Match:	73.98%	Indels:	0
DB:	6	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x AAD41202 (1-2275)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20

Db 1914 TGTAAAGCAGTGTGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGATGGAAGGACT 1973

Qy 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40

Db 1974 CACACTGGAGAGAAACCTTATGAATGTAAGCAGTGTGGAAAGCTTTTGGATGTGCCCTCG 2033

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60

Db 2034 CGACTTCAAAATGATGGAAGGACTCACACTGGAGAGAAACCTTATGAATGTAAGCAGTGT 2093

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77

Db 2094 GGGAAAGCTTTTGGATGTCCCTCAAAACCTTGAAGGATGGAAGGACTCAC 2144

RESULT 12

ADM03740

ID ADM03740 standard; cDNA; 2280 BP.

```

XX ADM03740;
XX AC
XX DT 20-MAY-2004 (first entry)
XX DE Human cDNA of the invention SEQ ID NO:2425.
XX KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX OS Homo sapiens.
XX PN EP1347046-A1.
XX PD 24-SEP-2003.
XX PF 12-APR-2002; 2002EP-00008400.
XX PR 22-MAR-2002; 2002JP-00137785.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI; 2003-723558/69.
XX DR P-PSDB; ADM06183.
XX PT New polynucleotides and polypeptides are useful in gene therapy, for
XX developing a diagnostic marker or medicines for regulating their
XX expression and activity, or as a target of gene therapy.
XX PS Claim 1; SEQ ID NO 2425; 305pp; English.
XX CC The invention relates to a novel human polynucleotide and the encoded
XX polypeptide. A polynucleotide of the invention may have a use in gene
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX as a primer for synthesizing the polynucleotide or as a probe for
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX useful in gene therapy, for developing a diagnostic marker or medicines
XX for regulating their expression and activity, or as a target of gene
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX are useful as pharmaceutical agents. The present sequence represents a
XX cDNA sequence of the invention.
XX SQ Sequence 2280 BP; 729 A; 470 C; 534 G; 547 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.3e-30 Length: 2280
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 11 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ADM03740 (1-2280)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1861 TGTAAACGACGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGTCATGGAAGGACT 1920
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1921 CACACTGGAGAACCCCTATGATGTAAGCAGTGTGGGAAGTCTTTTGGATGTGCTCG 1980
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 1981 CGACTTCAAAATGCTGAAGGACTCACACTGGAGAAACCGTATAAATGAAGCAATCT 2040
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2041 GGGAAAGCTTTTGGATGTCCCTCAACCTTCAAGGACGATGAAGGACTCAC 2091

```

```

RESULT 13
ADO00362
ID ADO00362 standard; cDNA; 2718 BP.
XX AC
XX AD ADO00362;
XX DT 29-JUL-2004 (first entry)
XX DE Novel human cDNA sequence #1177.
XX KW ds; gene; anti-inflammatory; dermatological; neuroprotective;
XX immunomodulator; antibacterial; virucide; antipsoriatic; cytostatic;
XX gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
XX psoriasis; diabetes; early aging; hormonal imbalance;
XX ischemic heart disease; ulcerative colitis.
XX OS Homo sapiens.
XX PN WO2004038003-A2.
XX PD 06-MAY-2004.
XX PF 24-OCT-2003; 2003WO-US033947.
XX PR 25-OCT-2002; 2002US-0421061P.
XX PR 25-OCT-2002; 2002US-0421080P.
XX PR 25-OCT-2002; 2002US-0421552P.
XX PR 25-OCT-2002; 2002US-0421614P.
XX PR 30-OCT-2002; 2002US-0422177P.
XX PR 30-OCT-2002; 2002US-0422178P.
XX PR 15-NOV-2002; 2002US-0426355P.
XX PR 15-NOV-2002; 2002US-0426384P.
XX PR 15-NOV-2002; 2002US-0426394P.
XX PR 15-NOV-2002; 2002US-0426430P.
XX PR 15-NOV-2002; 2002US-0426916P.
XX PR 27-NOV-2002; 2002US-0429224P.
XX PR 27-NOV-2002; 2002US-0429275P.
XX PR 27-NOV-2002; 2002US-0429302P.
XX PR 27-NOV-2002; 2002US-0429326P.
XX PR 27-NOV-2002; 2002US-0429651P.
XX PR 04-DEC-2002; 2002US-0430645P.
XX PR 04-DEC-2002; 2002US-0430651P.
XX PR 04-DEC-2002; 2002US-0430657P.
XX PR 04-DEC-2002; 2002US-0430663P.
XX PR 04-DEC-2002; 2002US-0430668P.
XX PR 04-DEC-2002; 2002US-0430684P.
XX PR 05-DEC-2002; 2002US-0430937P.
XX PR 05-DEC-2002; 2002US-0430965P.
XX PR 05-DEC-2002; 2002US-0431458P.
XX PR 12-DEC-2002; 2002US-0433251P.
XX PR 12-DEC-2002; 2002US-0433500P.
XX PR 13-DEC-2002; 2002US-0433316P.
XX PR 13-DEC-2002; 2002US-0433318P.
XX PR 23-DEC-2002; 2002US-0436236P.
XX PR 03-JAN-2003; 2003US-0437914P.
XX PR 17-JAN-2003; 2003US-0440820P.
XX PR 17-JAN-2003; 2003US-0440821P.
XX PR 18-APR-2003; 2003US-0463700P.
XX PR 18-APR-2003; 2003US-0463708P.
XX PR 18-APR-2003; 2003US-0463716P.
XX PR 18-APR-2003; 2003US-0463732P.
XX PR 02-MAY-2003; 2003US-0467199P.
XX PR 02-MAY-2003; 2003US-0467201P.
XX PR 02-MAY-2003; 2003US-0467203P.
XX PR 02-MAY-2003; 2003US-0467230P.
XX PR 19-MAY-2003; 2003US-0471306P.
XX PR 19-MAY-2003; 2003US-0471336P.
XX PR 22-MAY-2003; 2003US-0472420P.
XX PR 22-MAY-2003; 2003US-0472430P.
XX PR 09-JUN-2003; 2003US-0476609P.
XX PR 09-JUN-2003; 2003US-0476621P.
XX PR 09-JUN-2003; 2003US-0476632P.
XX PR 09-JUN-2003; 2003US-0476641P.

```

PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485329P.
PR 14-JUL-2003; 2003US-0485446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI; 2004-365511/34.
DR P-PSDB; ADN99577.
XX

PT New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX

PS Claim 1; SEQ ID NO 1961; 532pp; English.

XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.
XX

SQ Sequence 2718 BP; 970 A; 530 C; 612 G; 606 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.646-30 Length: 2718
Score: 327.00 Matches: 56
Percent Similarity: 81.8% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 12 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ADO00362 (1-2718)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1687 TGTAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAAATGCATGGAAGGACT 1746
QY 21 HisThrGlyGluLysProTyrLysCysGlyGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1747 CACACTGGAGAGAACCCCTATGATGTAAGCAGTGTGGGAAGTCTTTGGATGTGCCCTCG 1806
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1807 CGACTTCAAATGCATGGAAGGACTCACACTGGAGAGAACCCGTATATATGTAAGCAATGT 1866
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1867 GGGAAAGCTTTTGGATGTCCCTCAACCTTCGAGGCATGGAAGGACTCAC 1917

RESULT 14

ADN98793

ID ADN98793 standard; cDNA; 2718 BP.

XX

AC ADN98793;

XX

DT 29-JUL-2004 (first entry)
XX Novel human cDNA sequence #393.
DE
XX
KW ds; gene; anti-inflammatory; dermatological; neuroprotective;
KW immunomodulator; antibacterial; virucide; antipsoxiatic; cytostatic;
KW gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
KW psoriasis; diabetes; early aging; hormonal imbalance;
KW ischemic heart disease; ulcerative colitis.
XX
OS Homo sapiens.
XX
FN WO2004038003-A2.
XX
XX
PD 06-MAY-2004.
XX
XX 24-OCT-2003; 2003WO-US033947.
XX
XX 25-OCT-2002; 2002US-0421061P.
PR 25-OCT-2002; 2002US-0421080P.
PR 25-OCT-2002; 2002US-0421552P.
PR 25-OCT-2002; 2002US-0421614P.
PR 30-OCT-2002; 2002US-0422177P.
PR 30-OCT-2002; 2002US-0422178P.
PR 15-NOV-2002; 2002US-0426355P.
PR 15-NOV-2002; 2002US-0426384P.
PR 15-NOV-2002; 2002US-0426394P.
PR 15-NOV-2002; 2002US-0426430P.
PR 15-NOV-2002; 2002US-0426916P.
PR 27-NOV-2002; 2002US-0429224P.
PR 27-NOV-2002; 2002US-0429275P.
PR 27-NOV-2002; 2002US-0429302P.
PR 27-NOV-2002; 2002US-0429326P.
PR 27-NOV-2002; 2002US-0429651P.
PR 04-DEC-2002; 2002US-0430645P.
PR 04-DEC-2002; 2002US-0430651P.
PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430686P.
PR 04-DEC-2002; 2002US-0430684P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 12-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.

PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI: 2004-365511/34.
XX P-PSDB; ADN99577.

XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX

PS Claim 1; SEQ ID NO 393; 532pp; English.

XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.
XX

XX Sequence 2718 BP; 970 A; 530 C; 612 G; 606 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-64e-30 Length: 2718
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 12 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ADN98793 (1-2718)

QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
Db 1687 TGTAAAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 1746
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1747 CACACTGGAGAGAACCCCTATGAATGTAAGCAGTGTGGGAAGCTTTTGGATGTGCTCG 1806
QY 41 AsnLeuThrLysHisLysLysLeuHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 1807 CGACTTCAATCATGGAAGGACTCACACTGGAGAGAACCCCTATGAATGTAAGCAGTGT 1866
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1867 GGGAAAGCTTTTGGATGTCCCTCAACCTTCGAAGCATGGAAGGACTCAC 1917

RESULT 15

ADN99082
ID ADN99082 standard; cDNA; 2718 BP.

XX
XX AC
XX ADN99082;

XX DT 29-JUL-2004 (first entry)

XX DE Novel human cDNA sequence #682.

XX ds; gene; anti-inflammatory; dermatological; neuroprotective;
KW immunomodulator; antibacterial; virucide; antipsoriatic; cytostatic;

gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
psoriasis; diabetes; early aging; hormonal imbalance;
ischemic heart disease; ulcerative colitis.

Homo sapiens.

WO2004038003-A2.

06-MAY-2004.

24-OCT-2003; 2003WO-US033947.

25-OCT-2002; 2002US-0421061P.

25-OCT-2002; 2002US-0421080P.

25-OCT-2002; 2002US-0421552P.

25-OCT-2002; 2002US-0421614P.

30-OCT-2002; 2002US-0422177P.

30-OCT-2002; 2002US-0422178P.

15-NOV-2002; 2002US-0426355P.

15-NOV-2002; 2002US-0426384P.

15-NOV-2002; 2002US-0426394P.

15-NOV-2002; 2002US-0426430P.

15-NOV-2002; 2002US-0426916P.

27-NOV-2002; 2002US-0429224P.

27-NOV-2002; 2002US-0429275P.

27-NOV-2002; 2002US-0429302P.

27-NOV-2002; 2002US-0429326P.

04-DEC-2002; 2002US-0430651P.

04-DEC-2002; 2002US-0430651P.

04-DEC-2002; 2002US-0430657P.

04-DEC-2002; 2002US-0430663P.

04-DEC-2002; 2002US-0430668P.

05-DEC-2002; 2002US-0430684P.

05-DEC-2002; 2002US-0430937P.

05-DEC-2002; 2002US-0430965P.

12-DEC-2002; 2002US-0431458P.

12-DEC-2002; 2002US-043251P.

13-DEC-2002; 2002US-0433500P.

13-DEC-2002; 2002US-0433316P.

23-DEC-2002; 2002US-0433318P.

03-JAN-2003; 2002US-0436236P.

17-JAN-2003; 2003US-0437914P.

17-JAN-2003; 2003US-0440820P.

18-APR-2003; 2003US-0463700P.

18-APR-2003; 2003US-0463708P.

18-APR-2003; 2003US-0463716P.

02-MAY-2003; 2003US-0463732P.

02-MAY-2003; 2003US-0467199P.

02-MAY-2003; 2003US-0467201P.

02-MAY-2003; 2003US-0467203P.

19-MAY-2003; 2003US-0467230P.

19-MAY-2003; 2003US-0471336P.

22-MAY-2003; 2003US-0472420P.

22-MAY-2003; 2003US-0472430P.

09-JUN-2003; 2003US-0476609P.

09-JUN-2003; 2003US-0476621P.

09-JUN-2003; 2003US-0476632P.

09-JUN-2003; 2003US-0476641P.

08-JUL-2003; 2003US-0485217P.

08-JUL-2003; 2003US-0485218P.

08-JUL-2003; 2003US-0485223P.

08-JUL-2003; 2003US-0485224P.

08-JUL-2003; 2003US-0485325P.

14-JUL-2003; 2003US-0485359P.

14-JUL-2003; 2003US-0486446P.

14-JUL-2003; 2003US-0486480P.

15-JUL-2003; 2003US-0486891P.

08-AUG-2003; 2003US-0486960P.

08-AUG-2003; 2003US-0493341P.

08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
 PI Wong JGP, Wu G, Zhang H, Zeng C;
 XX WPI; 2004-365511/34.
 DR P-P8DB; ADN99866.
 XX
 PT New nucleic acid molecules, useful in preparing a composition for
 PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
 PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
 PT ulcerative colitis.
 XX
 PS Claim 1; SEQ ID NO 682; 532pp; English.
 XX
 CC The invention relates to a nucleic acid molecule comprising a
 CC polynucleotide sequence or its complement that encodes a polypeptide. The
 CC nucleic acid is useful in preparing a composition for treating or
 CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
 CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
 CC heart disease or ulcerative colitis. This sequence corresponds to a
 CC nucleic acid of the invention.
 XX
 SQ Sequence 2718 BP; 970 A; 530 C; 612 G; 606 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.64e-30 Length: 2718
 Score: 327.00 Matches: 56
 Percent Similarity: 81.82% Conservative: 7
 Best Local Similarity: 72.73% Mismatches: 14
 Query Match: 73.98% Indels: 0
 DB: 12 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x ADN99082 (1-2718)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 Db 1687 TGTAAAGCAGTGTGGGAAAGCCCTTTGGATCTGCCTCACACCTTCAAAATGCGAAGGACT 1746
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 1747 CACACTGGAGAGAAACCCCTATGAAATGTAGCAGTGTGGAAAGTCTTTGGATGTGCCCTCG 1806
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 Db 1807 CGACTTCAAAATGCGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAGCAATGT 1866
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 1867 GGGAAAGCTTTGGATGTCCCTCAAAACCTTCGAAGGCATGGAAGGACTCAC 1917

Search completed: January 12, 2006, 02:38:53
 Job time : 467 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 01:29:19 ; Search time 155 Seconds

(without alignments)
883.047 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCGAFIQKSLIRHQT.....KQCGAFGCPNLRHGRTH 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US1069861/runat_09012006_144453_20393/app_query.fasta_1.263
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1069861@cgn1.1.193@runat_09012006_144453_20393 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCUTS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	315	71.3	601	3	US-09-949-016-27361 Sequence 27361, A
c 2	315	71.3	601	3	US-09-949-016-149636 Sequence 149636,
c 3	315	71.3	1914	3	US-10-104-047-1464 Sequence 1464, Ap
4	315	71.3	3798	3	US-09-949-016-4204 Sequence 4204, Ap
5	315	71.3	3839	3	US-09-949-016-485 Sequence 485, App
6	315	71.3	156942	3	US-09-949-016-12227 Sequence 12227, A
7	315	71.3	156950	3	US-09-949-016-15946 Sequence 15946, A
8	314	71.0	3186	3	US-09-016-434-1390 Sequence 1390, Ap
9	313	70.8	1835	3	US-10-104-047-901 Sequence 901, App

10	309	69.9	601	3	US-09-949-016-27493 Sequence 27493, A
11	309	69.9	601	3	US-09-949-016-92521 Sequence 92521, A
12	309	69.9	1694	3	US-10-104-047-1500 Sequence 1500, Ap
13	309	69.9	1820	3	US-09-949-016-2645 Sequence 2645, Ap
14	309	69.9	1839	3	US-09-949-016-493 Sequence 493, App
15	309	69.9	2241	3	US-10-104-047-693 Sequence 693, App
16	309	69.9	23174	3	US-09-949-016-14387 Sequence 14387, A
17	309	69.9	23187	3	US-09-949-016-12235 Sequence 12235, A
18	308	69.7	2110	3	US-10-104-047-1778 Sequence 1778, Ap
19	307	69.5	2320	3	US-09-016-434-1054 Sequence 1054, Ap
20	307	69.5	4272	3	US-09-620-312D-586 Sequence 586, App
21	306	69.2	601	3	US-09-949-016-44799 Sequence 44799, A
22	306	69.2	2197	3	US-10-104-047-884 Sequence 884, App
23	305	69.0	2038	3	US-09-620-312D-911 Sequence 911, App
24	305	69.0	3078	3	US-10-104-047-622 Sequence 622, App
25	305	69.0	45138	3	US-09-949-016-13027 Sequence 13027, A
c 26	304	68.8	563	3	US-09-385-982-433 Sequence 433, App
27	302	68.3	1429	3	US-09-023-655-689 Sequence 689, App
28	302	68.3	1995	3	US-10-104-047-1730 Sequence 1730, Ap
29	301	68.1	2210	3	US-10-104-047-1453 Sequence 1453, Ap
30	301	68.1	3138	3	US-10-104-047-83 Sequence 83, Appl
31	299	67.6	2351	3	US-09-016-434-1337 Sequence 1337, Ap
32	299	67.6	2637	3	US-09-949-016-5823 Sequence 5823, Ap
33	299	67.6	3754	3	US-10-104-047-530 Sequence 530, App
34	299	67.6	4468	3	US-09-620-312D-243 Sequence 243, App
c 35	298	67.4	601	3	US-09-949-016-17365 Sequence 17365, A
36	298	67.4	601	3	US-09-949-016-27359 Sequence 27359, A
c 37	298	67.4	601	3	US-09-949-016-27360 Sequence 27360, A
c 38	298	67.4	601	3	US-09-949-016-149637 Sequence 149637, A
c 39	298	67.4	601	3	US-09-949-016-149638 Sequence 149638, A
40	298	67.4	2441	3	US-09-949-016-2756 Sequence 2756, Ap
41	298	67.4	19861	3	US-09-949-016-14498 Sequence 14498, A
42	297	67.2	2487	3	US-10-104-047-898 Sequence 898, App
43	295	66.7	2779	3	US-09-976-594-175 Sequence 175, App
44	294	66.5	1460	3	US-09-774-528-413 Sequence 413, App
45	294	66.5	1460	3	US-10-120-988-413 Sequence 413, App

ALIGNMENTS

RESULT 1
US-09-949-016-27361/c
; Sequence 27361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27361
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27361

Alignment Scores:
Pred. No.: 2.74e-31 Length: 601
Score: 315.00 Matches: 56
Percent Similarity: 80.52% Conservative: 6
Best Local Similarity: 72.73% Mismatches: 15
Query Match: 71.27% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-27361 (1-601)

```
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 274 TGTAAAGAAATGGCAAGCAATTTATATGGTCTTCAACCCCTAACTAGACATAAGAGGATA 215
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 214 CACACTGGAGAGAGCCCTACAAATGTGAAGATGTGGCAAGCTTTTAGCGGTCTCTCA 155
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 154 ACCCTTACTAGCATAAGACAATTCATCTCGAGAGAAACCCCTACAAATGTAAAGAATGT 95
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 94 GGCAAGAGCTTTTAAGCACTCTCTCAGCCCTTGCTAAACATAAAATAATACAT 44
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 2

US-09-949-016-149636/c
; Sequence 149636, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 149636

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-149636

Alignment Scores:

Pred. No.:	2,748-31	Length:	601
Score:	315.00	Matches:	56
Percent Similarity:	80.52%	Conservative:	6
Best Local Similarity:	72.73%	Mismatches:	15
Query Match:	71.27%	Indels:	0
DB:	3	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-149636 (1-601)

```
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 274 TGTAAAGAAATGGCAAGCAATTTATATGGTCTTCAACCCCTAACTAGACATAAGAGGATA 215
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 214 CACACTGGAGAGAGCCCTACAAATGTGAAGATGTGGCAAGCTTTTAGCGGTCTCTCA 155
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 154 ACCCTTACTAGCATAAGACAATTCATCTCGAGAGAAACCCCTACAAATGTAAAGAATGT 95
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 94 GGCAAGAGCTTTTAAGCACTCTCTCAGCCCTTGCTAAACATAAAATAATACAT 44
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 3

US-10-104-047-1464
; Sequence 1464, Application US/10104047
; Patent No. 6943241

; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: No. 6943241el full length cDNA

; FILE REFERENCE: HI-A0105

; CURRENT APPLICATION NUMBER: US/10/104,047

; CURRENT FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER:

; PRIOR FILING DATE:

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1464

; LENGTH: 1914

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-104-047-1464

Alignment Scores:

Pred. No.:	1,35e-30	Length:	1914
Score:	315.00	Matches:	55
Percent Similarity:	77.92%	Conservative:	5
Best Local Similarity:	71.43%	Mismatches:	17
Query Match:	71.27%	Indels:	0
DB:	3	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-104-047-1464 (1-1914)

```
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1557 TGCAGTGAATGTGGAAAGCTTTTATTTCAGATGTCAAACCTCATTCGACACAGAGAATT 1616
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 21 HisThrGlyLysGlyProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1617 CATACGGGTGAGAAACCCCTATGCATGTACATGTGTGAAAAAGCCTTTAGTCAGAAAAATCA 1676
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1677 AACCTCACTGAACATGAGAAATTCATCTCGAGAGAAACCTTATCATTTGTAATCAATGT 1736
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1737 GGGAAAGCTTTTCAGTCAGAGACAAAATCTTCTTGAGCATGAAAAAATTCAT 1787
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 4

US-09-949-016-4204
; Sequence 4204, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4204

; LENGTH: 3798

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-4204

Alignment Scores:

Pred. No.:	3.49e-30	Length:	3798
Score:	315.00	Matches:	56
Percent Similarity:	80.52%	Conservative:	6
Best Local Similarity:	72.73%	Mismatches:	15
Query Match:	71.27%	Indels:	0


```

DB:                                     3           Gaps:                               0
US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-4204 (1-3798)
QY      1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2368 TGTAAAGAATGTGGCAAGCATTTATATGCTTCAACCTTACTAGACATAAGAGGATA 2427
QY      21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2428 CACACTGGAGAGAGCCCTACAATGTGAAGATGTGGCAAGCTTTTAGCCGTTCTCTCA 2487
QY      41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2488 ACCCTTACTAAGCATAAGACAATTCATCTGGAGAGAAACCCCTACAAATGTAAGAATGT 2547
QY      61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2548 GGCAAGCTTTTAAGCACTCTCAGCCCTTGCTAAACATAAAATAATACAT 2598

RESULT 5
US-09-949-016-485
; Sequence 485, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 485
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(156942)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12227

Alignment Scores:
Pred. No.:          5,948-28          Length:          156942
Score:              315.00           Matches:         56
Percent Similarity: 80.52%           Conservatives:    6
Best Local Similarity: 72.73%         Mismatches:      15
Query Match:        71.27%           Indels:          0
DB:                  3                Gaps:            0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-12227 (1-156942)
QY      1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153519 TGTAAAGAATGTGGCAAGCATTTATATGCTTCAACCTTACTAGACATAAGAGGATA 153578
QY      21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153579 CACACTGGAGAGAGCCCTACAATGTGAAGATGTGGCAAGCTTTTAGCCGTTCTCTCA 153638
QY      41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153639 ACCCTTACTAAGCATAAGACAATTCATCTGGAGAGAAACCCCTACAAATGTAAGAATGT 153698
QY      61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153699 GGCAAGCTTTTAAGCACTCTCAGCCCTTGCTAAACATAAAATAATACAT 153749

RESULT 7
US-09-949-016-15946
; Sequence 15946, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15946

```

```

; Sequence 12227, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12227
; TYPE: DNA
; LENGTH: 156942
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(156942)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12227

Alignment Scores:
Pred. No.:          5,948-28          Length:          156942
Score:              315.00           Matches:         56
Percent Similarity: 80.52%           Conservatives:    6
Best Local Similarity: 72.73%         Mismatches:      15
Query Match:        71.27%           Indels:          0
DB:                  3                Gaps:            0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-12227 (1-156942)
QY      1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153519 TGTAAAGAATGTGGCAAGCATTTATATGCTTCAACCTTACTAGACATAAGAGGATA 153578
QY      21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153579 CACACTGGAGAGAGCCCTACAATGTGAAGATGTGGCAAGCTTTTAGCCGTTCTCTCA 153638
QY      41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153639 ACCCTTACTAAGCATAAGACAATTCATCTGGAGAGAAACCCCTACAAATGTAAGAATGT 153698
QY      61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      153699 GGCAAGCTTTTAAGCACTCTCAGCCCTTGCTAAACATAAAATAATACAT 153749

RESULT 7
US-09-949-016-15946
; Sequence 15946, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15946

```



```

QY 41 AsnLeuThrLysHisLysLysLysLeuHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 813 GAACCTTACCGCATCAGAGAGCTCATCTGCTGAAAGAACCTTATGAGTGAAGGAATGT 872

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 873 GGAAAGCCCTTACTTGTCAGCAGAACTTGTTCGACATCAAAAAAGTTCCAC 923

RESULT 10
US-09-949-016-27493
; Sequence 27493, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27493
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27493

Alignment Scores:
Pred. No.: 1,61e-30 Length: 601
Score: 309.00 Matches: 51
Percent Similarity: 84.42% Conservative: 14
Best Local Similarity: 66.23% Mismatches: 12
Query Match: 69.91% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-27493 (1-601)
QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 23 TGCAGTGAAGTGTGGAAAGCCCTTCATTTTCAGTCTTCCCTTAAGAAACATGAGATCT 82

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 83 CATACTGGAGAGAGCCCTTATGAGTGTGATCACTGTGGAAATCCTTTAGCCAGAGCTCT 142

QY 41 AsnLeuThrLysHisLysLysLysLeuHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 143 CATCTGAATGTGCACAAAAGAACTCACACTGGAGAGAAACCTTATGACTGTAAAGGAATGT 202

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 203 GGAAGGCTTTCAGTGTCTTCCTTCATCTCCCTTCAGAAACATGTGAGAACCCAC 253

RESULT 11
US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-27493 (1-601)
QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 23 TGCAGTGAAGTGTGGAAAGCCCTTCATTTTCAGTCTTCCCTTAAGAAACATGAGATCT 82

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 83 CATACTGGAGAGAGCCCTTATGAGTGTGATCACTGTGGAAATCCTTTAGCCAGAGCTCT 142

QY 41 AsnLeuThrLysHisLysLysLysLeuHisThrGlyGluLysProTyrLysCysLysGlnCys 60
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 143 CATCTGAATGTGCACAAAAGAACTCACACTGGAGAGAAACCTTATGACTGTAAAGGAATGT 202

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 203 GGAAGGCTTTCAGTGTCTTCCTTCATCTCCCTTCAGAAACATGTGAGAACCCAC 253

RESULT 12
US-10-104-047-1500
; Sequence 1500, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1500
; LENGTH: 1694
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1500

Alignment Scores:
Pred. No.: 6,74e-30 Length: 1694
Score: 309.00 Matches: 52
Percent Similarity: 80.52% Conservative: 10
Best Local Similarity: 67.53% Mismatches: 15
Query Match: 69.91% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-104-047-1500 (1-1694)
QY 1 CysLysAspCysGlyLysAlaPheLeuGlnLysSerAsnLeuLeuArgHisGlnArgThr 20
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1192 TGTAAAGGAATGTGGCAAGCCCTTTACCAGAGCTTCAAGCTTGTTCACATCAGAGATTT 1251

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```
Db 1252 CATAGTGTGAGAAACCCATGAGTGCAGGAATGTGGCAAGCCTTTAGTAGTGGCTCA 1311
      |||
Qy 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysLysGlnCys 60
      |||
Db 1312 GCACCTACTAATCATCATCAGAGAATTCACACTGCTGAGAAACCCCTATGATTGTAAGGAATGT 1371
      |||
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
      |||
Db 1372 GGAAGGCTTTTACTCAGAGCTCAGAGCTTCGTCAACATCATCAGAGAATTCAC 1422
      |||

RESULT 13
US-09-949-016-2645
; Sequence 2645, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2645
; LENGTH: 1820
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2645

Alignment Scores:
Pred. No.: 7,44e-30 Length: 1820
Score: 309.00 Matches: 51
Percent Similarity: 84.42% Conservatve: 14
Best Local Similarity: 66.23% Mismatches: 12
Query Match: 69.91% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-2645 (1-1820)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
      |||
Db 817 TGCAGTGACTGTGGAAAGCCTTCATTTTCAGTCTTCCCTTAAGAAACACATGAGATCT 876
      |||
Qy 21 HisthrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
      |||
Db 877 CATACTGAGAGAAGCCTTATGAGTGTGATCACTGTGGAAATCCTTTAGCCAGAGCTCT 936
      |||
Qy 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysLysGlnCys 60
      |||
Db 937 CATCTGATGTGCACAAAGAACTCACACTGGAGAAACCCCTATGACTGTAAGGAATGT 996
      |||
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
      |||
Db 997 GGAAGGCTTTTACTGTTCCCTTCATCCTTCAGAAACATGTGAGAAACCCAC 1047
      |||

RESULT 14
US-09-949-016-493
; Sequence 493, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-493

Alignment Scores:
Pred. No.: 7,55e-30 Length: 1839
Score: 309.00 Matches: 51
Percent Similarity: 84.42% Conservatve: 14
Best Local Similarity: 66.23% Mismatches: 12
Query Match: 69.91% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-949-016-493 (1-1839)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
      |||
Db 832 TGCAGTGACTGTGGAAAGCCTTCATTTTCAGTCTTCCCTTAAGAAACACATGAGATCT 891
      |||
Qy 21 HisthrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
      |||
Db 892 CATACTGAGAGAAGCCTTATGAGTGTGATCACTGTGGAAATCCTTTAGCCAGAGCTCT 951
      |||
Qy 41 AsnLeuThrLysHisLysLysLysHisThrGlyGluLysProTyrLysCysLysGlnCys 60
      |||
Db 952 CATCTGAATGTGCACAAAGAACTCACACTGGAGAAACCCCTATGACTGTAAGGAATGT 1011
      |||
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
      |||
Db 1012 GGAAGGCTTTTACTGTTCCCTTCATCCTTCAGAAACATGTGAGAAACCCAC 1062
      |||


RESULT 15
US-10-104-047-693
; Sequence 693, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 693
; LENGTH: 2241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-693

Alignment Scores:
Pred. No.: 9,91e-30 Length: 2241
Score: 309.00 Matches: 52
Percent Similarity: 80.52% Conservatve: 10
Best Local Similarity: 67.53% Mismatches: 15
Query Match: 69.91% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-104-047-693 (1-2241)
Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
      |||
Db 1753 TGTAAAGGAATGTGGCAAGCCTTTTACCAGAGCTCAAAGCTTGTTCACATCATCAGAGAAT 1812
      |||
```

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 |||||
 Db 1813 CATACTGGTGAGAAACCCCTATGAGTGCAGGAATGTGGCAAGCCCTTTAGTAGTGGCTCA 1872
 |||||
 QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 |||||
 Db 1873 GCACCTTACTAATCATCAGAGAAATTCACACTGGGTGAGAAACCCCTATGATTGTAAGGAATGT 1932
 |||||
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 |||||
 Db 1933 GGAAGGCTTTTACTCAGAGCTCAGAGCTTCGTCAACATCAGAGAAATTCAC 1983
 |||||

Search completed: January 12, 2006, 03:42:50
 Job time : 182 secs

 This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 01:19:30 ; Search time 614 Seconds
(without alignments)

1037.039 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence:

1 CKDCGKAFQKSLNRHRT.....KCGKAFGCPNLRHRGRTH 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19597084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10669861/runat_09012006_144454_20413/app_query.fasta_1.263
-DB=Published Applications NA Main -OFFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEVTIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	1137	6	US-10-314-669-1
2	442	100.0	1137	8	US-10-669-861-1
3	374	84.6	630	6	US-10-314-669-13
4	374	84.6	630	8	US-10-669-861-13
5	327	74.0	1668	9	US-10-450-763-22733
6	327	74.0	1668	7	US-10-466-164-14
7	327	74.0	2280	6	US-10-108-260A-2425
8	326	73.8	636	6	US-10-314-669-263

9	325	73.5	3300	3	US-09-765-555-18
10	321	72.6	330	3	US-09-765-555-64
11	321	72.6	3300	3	US-09-765-555-14
12	321	72.6	3300	3	US-09-765-555-15
13	319	72.2	2175	9	US-10-450-763-18377
14	319	72.2	2667	6	US-10-108-260A-1226
15	318	71.9	505	6	US-10-029-386-2047
16	316	71.5	554	4	US-09-925-065A-189236
17	316	71.5	2253	10	US-11-037-295-22
18	315	71.3	468	3	US-09-918-995-3107
19	315	71.3	630	6	US-10-314-669-9
20	315	71.3	630	8	US-10-669-861-9
21	315	71.3	632	4	US-09-925-065A-488841
22	315	71.3	1258	6	US-10-029-386-25281
23	315	71.3	1914	6	US-10-104-047-1464
24	315	71.3	2664	6	US-10-108-260A-511
25	315	71.3	2690	8	US-10-723-860-8039
26	315	71.3	3639	9	US-10-450-763-390
27	315	71.3	3839	6	US-10-172-118-910
28	315	71.3	3839	7	US-10-342-887-910
29	315	71.3	3839	9	US-10-956-157-1072
30	314	71.0	283	6	US-10-029-386-26998
31	314	71.0	590	6	US-10-029-386-13298
32	314	71.0	3186	5	US-10-171-581-141
33	314	71.0	3186	6	US-10-305-720-1390
34	314	71.0	3213	5	US-10-098-841-141
35	314	71.0	3382	3	US-09-971-392-83
36	313	70.8	1835	6	US-10-104-047-901
37	312	70.6	392	3	US-09-864-761-20124
38	311	70.4	1945	6	US-10-094-749-1120
39	311	70.4	2662	7	US-10-363-616-66
40	310	70.1	630	6	US-10-314-669-259
41	310	70.1	630	8	US-10-669-861-259
42	310	70.1	2873	7	US-10-755-889-369
43	310	70.1	3300	3	US-09-765-555-17
44	309	69.9	514	6	US-10-029-386-9811
45	309	69.9	1694	6	US-10-104-047-1500

ALIGNMENTS

RESULT 1

US-10-314-669-1
; Sequence 1, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12779-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
;
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1134)
US-10-314-669-1

Alignment Scores:
Pred. No.: 9,72e-52 Length: 1137
Score: 442.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-314-669-1 (1-1137)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 97 TGTAAAGATTGGGGAAAGCTTTTCATTTCAGAAAGTCAAAACCTTCATCAGACACAGAGAACT 156

Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 157 CACACCGGGGAAAAACCGTACAGGTGTGAAGATGTGGCAAGCTTTTACCCTCAATCCTCA 216

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 217 AACCTTACTAAACATTAAGAAATTCATACCGGGGAAAAACCGTATAAATGTAAAGCAATGT 276

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 277 GGGAAGCTTTTGGATGTCCTCAAACTTCGAAGGCATGGAAGGACTCAC 327

RESULT 2
US-10-669-861-1
; Sequence 1, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Lee, Yangsoon
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; NAME/KEY: CDS
; LOCATION: (1)...(1134)
US-10-669-861-1

Alignment Scores:
Pred. No.: 9,72e-52 Length: 1137
Score: 442.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 97 TGTAAAGATTGGGGAAAGCTTTTCATTTCAGAAAGTCAAAACCTTCATCAGACACAGAGAACT 156

Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 157 CACACCGGGGAAAAACCGTACAGGTGTGAAGATGTGGCAAGCTTTTACCCTCAATCCTCA 216

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 217 AACCTTACTAAACATTAAGAAATTCATACCGGGGAAAAACCGTATAAATGTAAAGCAATGT 276

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 277 GGGAAGCTTTTGGATGTCCTCAAACTTCGAAGGCATGGAAGGACTCAC 327

RESULT 3
US-10-314-669-13
; Sequence 13, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; NAME/KEY: CDS
; LOCATION: (1)...(627)
US-10-314-669-13

Alignment Scores:
Pred. No.: 1,67e-42 Length: 630
Score: 374.00 Matches: 61
Percent Similarity: 92.21% Conservative: 10
Best Local Similarity: 79.22% Mismatches: 6
Query Match: 84.62% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-314-669-13 (1-630)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 181 TGCAGTGAAGTGTGGCGGAGGCTTCAGCCAGAAAGTCAAAACCTTCATCATACACAGAGGACA 240
```


RESULT 6
US-10-466-164-14
; Sequence 14, Application US/10466164
; Publication No. US20040058365A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, SCOTT R.;
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOUR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy P.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAFFO, Abel;
; APPLICANT: MARWAHA, Rakesh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN, Jr., Edward H.;
; APPLICANT: PERALTA, Careyana H.; DAVID, Marie H.;
; APPLICANT: LEWIS, Samantha A.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1215 PCT

```
; CURRENT APPLICATION NUMBER: US/10/466,164
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: PCT/US02/01008
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/263,065
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,326
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,622
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 2192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058365A1 LI:1175131.1:2001JAN12
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 64, 81
; OTHER INFORMATION: a, t, c, g, or other
US-10-466-164-14

Alignment Scores:
Pred. No.: 3,54e-35 Length: 2192
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-466-164-14 (1-2192)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1845 TGTAAAGCAGTGTGGGAAAGCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 1904

Qy 21 HistThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1905 CACACTGGAGAGAAACCCCTATGAATGTAAAGCAGTGTGGGAAAGTCTTTTGGATGTCCTCG 1964

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1965 CGACTTCAATGCATGGAGGACTCACACTGGAGAGAAACCCGTATAATGTAAAGCAATGT 2024

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2025 GGGAAAGCTTTTGGATGTCCTCAAAACCTTCGAAGGCATGGAAGGACTCAC 2075

RESULT 7
US-10-108-260A-2425
; Sequence 2425, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2425
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2425

Alignment Scores:
Pred. No.: 3,73e-35 Length: 2280
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-108-260A-2425 (1-2280)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1861 TGTAAAGCAGTGTGGGAAAGCCTTTGGATCTGCCTCACACCTTCAAAATGCATGGAAGGACT 1920

Qy 21 HistThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1921 CACACTGGAGAGAAACCCCTATGAATGTAAAGCAGTGTGGGAAAGTCTTTTGGATGTCCTCG 1980

Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1981 CGACTTCAATGCATGGAGGACTCACACTGGAGAGAAACCCGTATAATGTAAAGCAATGT 2040

Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 2041 GGGAAAGCTTTTGGATGTCCTCAAAACCTTCGAAGGCATGGAAGGACTCAC 2091

RESULT 8
US-10-314-669-263
; Sequence 263, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(633)
US-10-314-669-263

Alignment Scores:
Pred. No.: 9,61e-36 Length: 636
```

Score: 326.00 Matches: 55
Percent Similarity: 85.71% Conservative: 11
Best Local Similarity: 71.43% Mismatches: 11
Query Match: 73.76% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-314-669-263 (1-636)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 97 TGTAAGATGCGGGAAGCTTCATTCAGAGTCAAACTCATCAGACCCAGAGACT 156
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 157 CACACCGGGGAAAACCGTATAGTGCCTGATGTTGGGAAGAGTTTATGTCAGATTCC 216
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
DB 217 AGCCTCATTCGCCACCGAGCGGACACACCGGGGAAAACCGTATGAGTGTCCAGATTGC 276
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 277 GGAAAGTCTTTAGGCAGAGCACCACCTCACTCGGACCGGAGGATCCAC 327

RESULT 9

US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPp3
US-09-765-555-18

Alignment Scores:
Pred. No.: 1,166-34 Length: 3300
Score: 325.00 Matches: 54
Percent Similarity: 84.42% Conservative: 11
Best Local Similarity: 70.13% Mismatches: 12
Query Match: 73.53% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-765-555-18 (1-3300)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 2758 TGTCGGAATGTTGGTAAGCTTCAGCCAGCAGCTCCCTGGTGGCCACCGGTACC 2817
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 2818 CACACGGGTGAAAACCGTATAGTGCCTGATGTTGGGAAGAGTTTATGTCAGATTCC 2877
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
DB 2878 AACCTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGTCAGAAATGT 2937
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77

DB 2938 GGCAAGTCTTTAGCCAGTCCAGCACTGGTGGCCACCAAGTACTCAC 2988
RESULT 10
US-09-765-555-64
; Sequence 64, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZFP2b
US-09-765-555-64

Alignment Scores:
Pred. No.: 2,056-35 Length: 330
Score: 321.00 Matches: 53
Percent Similarity: 83.12% Conservative: 11
Best Local Similarity: 68.83% Mismatches: 13
Query Match: 72.62% Indels: 0
DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-09-765-555-64 (1-330)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
DB 55 TGTCGGAATGTTGGTAAGCTTCCTCTCAGAGCTCTCACCTGGTGGCCACCGGTACC 114
QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
DB 115 CACACGGGTGAAAACCGTATAGTGCCTGATGTTGGGAAGAGTTTATGTCAGATTCC 174
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
DB 175 AACCTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGTCAGAAATGT 234
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
DB 235 GGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGCGAGCCACCAAGTACTCAC 285

RESULT 11
US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300

Cy	1	CysLysAspCysGlyLysAlaPheIleLysSerAsnLeuLeuArgHisGlnArgThr	20
		:::::	
		:::::	
Db	2758	TGTCCGAAGTGGTAGTCCTCTCTCAGAGCTCTCACTGTGTGCCACCGCGTACC	2817

US-10-108-260A-1226
; Sequence 1226, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
: APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
 ; FILE REFERENCE: HI-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108,260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1226
 ; LENGTH: 2667
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-108-260A-1226

Alignment Scores:
 Pred. No.: 6,12e-34 Length: 2667
 Score: 319.00 Matches: 54
 Percent Similarity: 84.42% Conservatives: 11
 Best Local Similarity: 70.13% Mismatches: 12
 Query Match: 72.17% Indels: 0
 DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-108-260A-1226 (1-2667)

QY 1 CysLysAspCysGlyLysAlaPheLleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
 Db 1412 TGTGAAGAATGTGGCAAGCTTTAAACCACTCTCATCCCTTACTAAACATAAGAAATTT 1471
 QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 Db 1472 CATACTGGAGAGAAACCCCTACAAATGTGAAGATGTGGCAAGCTTTTAAACCACTCTCT 1531
 QY 41 AsnLeuThrLysHisLysLysLysLeHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 Db 1532 TCCCTTACTAAACATAAGAAATTTACTACTGGAGAGAAACCCCTACAAATGTGAAGATCT 1591
 QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 1592 GGCAAGCTTTTAAACCACTCTCAACCTTATTAAACATAAGAAATTCAT 1642

RESULT 15
 US-10-029-386-2047/c
 ; Sequence 2047, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
 ; FILE REFERENCE: AEOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 2047
 ; LENGTH: 505
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO CHR19.1
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.7
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.99
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
 ; OTHER INFORMATION: EST HUMAN HIT: AW236498.1, EVALUE 0.00e+00
 ; OTHER INFORMATION: NT HIT: g114760087, EVALUE 5.00e-35
 ; OTHER INFORMATION: SWISSPROT HIT: P17024, EVALUE 6.00e-44
 US-10-029-386-2047

Alignment Scores:
 Pred. No.: 9.49e-35 Length: 505

Score: 318.00 Matches: 55
 Percent Similarity: 81.58% Conservatives: 7
 Best Local Similarity: 72.37% Mismatches: 14
 Query Match: 71.95% Indels: 0
 DB: 6 Gaps: 0
 US-10-669-861-2_COPY_33_109 (1-77) x US-10-029-386-2047 (1-505)
 QY 2 LysAspCysGlyLysAlaPheLleGlnLysSerAsnLeuIleArgHisGlnArgThrHis 21
 Db 503 AAGCAGTGTGGGAAGCCCTTTGGATCTGCCTCACACCTTCAATGTCATGGAAGACTCAC 444
 QY 22 ThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSerAsn 41
 Db 443 ACTGGAGAGAAACCCCTATGAATGTAAGCAGTGTGGGAAGTCTTTTGGATGTGCTCGCGA 384
 QY 42 LeuThrLysHisLysLysLysLeHisThrGlyGluLysProTyrLysCysLysGlnCysGly 61
 Db 383 CTTCAATGTCATGGAAGGACTCACACTGGAGAGAAACCGTATAAATGTAAAGCAATGTGGG 324
 QY 62 LysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
 Db 323 AAAGCTTTTGGATGTCCCTCAAAACCTTCGAAGGCATGGAAGGACTCAC 276

Search completed: January 12, 2006, 03:39:45
 Job time : 616 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 02:31:44 ; Search time 226 Seconds

(without alignments)
275.751 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCKAFIQKSLNRHQRT.....KQCKAFGCPNLRHRGTH 77

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10669861/runat_09012006_144454_20431/app_query.fasta_1.263
-DB=Published Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10669861@cgn 1.1.122 @runat_09012006_144454_20431
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Published Applications_NA_New.*
- 2: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	302	68.3	982	6	US-10-750-185-60532 Sequence 60532, A
C 2	302	68.3	982	6	US-10-750-623-54987 Sequence 54987, A
C 3	302	68.3	2196	7	US-11-000-688-253 Sequence 253, App
4	301	68.1	2124	7	US-11-136-527-713 Sequence 713, App
5	299	67.6	600	7	US-11-136-527-4391 Sequence 4391, App
6	299	67.6	1940	7	US-11-136-527-295 Sequence 295, App
7	298	67.4	600	7	US-11-136-527-6532 Sequence 6532, App
8	298	67.4	1578	7	US-11-136-527-2436 Sequence 2436, App

ALIGNMENTS

RESULT 1

US-10-750-185-60532/c
; Sequence 60532, Application US/10750185

; Publication No, US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIN version 3.1

; SEQ ID NO 60532

; LENGTH: 982

; TYPE: DNA

; ORGANISM: Bovine

US-10-750-185-60532

Alignment Scores:

Pred. No.:	1,15e-28	Length:	982
Score:	302.00	Matches:	50
Percent Similarity:	81.82%	Conservative:	13
Best Local Similarity:	64.94%	Mismatches:	14
Query Match:	68.33%	Indels:	0

Sequence 54987, A
Sequence 54987, A
Sequence 44166, A
Sequence 44166, A
Sequence 6854, App
Sequence 2758, App
Sequence 41858, A
Sequence 41858, A
Sequence 3886, App
Sequence 601, App
Sequence 129, App
Sequence 1179, App
Sequence 1046, App
Sequence 5142, App
Sequence 710, App
Sequence 28, Appl
Sequence 119, App
Sequence 55, Appl
Sequence 42613, A
Sequence 42613, A
Sequence 30010, A
Sequence 30010, A
Sequence 49223, A
Sequence 49223, A
Sequence 928, App
Sequence 61337, A
Sequence 61337, A
Sequence 58156, A
Sequence 58156, A
Sequence 138, App
Sequence 1136, App
Sequence 2585, App
Sequence 6227, App
Sequence 49213, A
Sequence 49213, A
Sequence 5181, App

```
DB:
US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-185-60532 (1-982)
QY 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 477 TGTAAAGAGTGTAAACAAAGCCTTTATTACAGCGCTCACATCTTACTATAACATCAGAAAAATT 418
QY 21 HisThrGlyLysGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 417 CATACTGGAGAGAGACCTTTATAAATGTAAAGAGTGTAAACAAAGCCTTTATTACAGCGCGCA 358
QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 357 AATCTTACTAAACATCAGCGAATTCATACGGGAGAGACCTTTATAAATGTAAAGAGTGT 298
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 297 AACAAAGCCTTTAATCAAGCTCAAGTCTTAACCTATCATCAGCGAATTCAT 247

RESULT 2
US-10-750-623-60532/c
; Sequence 60532, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: WMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60532
; LENGTH: 982
; TYPE: DNA
; ORGANISM: Bovine 19866881803796
US-10-750-623-60532

Alignment Scores:
Pred. No.: 1.15e-28 Length: 982
Score: 302.00 Matches: 50
Percent Similarity: 81.82% Conservative: 13
Best Local Similarity: 64.94% Mismatches: 14
Query Match: 68.33% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-60532 (1-982)
QY 1 CysLysAspCysGlyAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 477 TGTAAAGAGTGTAAACAAAGCCTTTATTACAGCGCTCACATCTTACTATAACATCAGAAAAATT 418
QY 21 HisThrGlyLysGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 417 CATACTGGAGAGAGACCTTTATAAATGTAAAGAGTGTAAACAAAGCCTTTATTACAGCGCGCA 358
QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 357 AATCTTACTAAACATCAGCGAATTCATACGGGAGAGACCTTTATAAATGTAAAGAGTGT 298
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 297 AACAAAGCCTTTAATCAAGCTCAAGTCTTAACCTATCATCAGCGAATTCAT 247

RESULT 3
US-10-750-623-60532/c
; Sequence 60532, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: WMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60532
; LENGTH: 982
; TYPE: DNA
; ORGANISM: Bovine 19866881803796
US-10-750-623-60532

Alignment Scores:
Pred. No.: 1.15e-28 Length: 982
Score: 302.00 Matches: 50
Percent Similarity: 81.82% Conservative: 13
Best Local Similarity: 64.94% Mismatches: 14
Query Match: 68.33% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-000-688-253 (1-2196)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1476 TGTAAAGACTGTGGGAAGCGCTTTTATTGTGGTTCAAGCCTTCATTCAGCATAAAGAATT 1535
QY 21 HisThrGlyLysGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1536 CACACAGGTGAGAAACCTATGAATGTCAAGAAATGTGGGAAGCGCTTCATTCAGTCAAT 1595
QY 41 AsnLeuThrLysHisLysLysLysHisThrGlyLysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1596 TACCTTACTCAGCATCAGAAAGATCCACACCGGTGAGAAGCGCTTCAGAAATGTAAAGAGTGT 1655
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1656 GGGAAAGCGCTTTTCGCTGGGGTTCGAGCCTCGTTAAGCACCAGAGGATACAT 1706

RESULT 4
US-11-136-527-713
; Sequence 713, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 713
; LENGTH: 2124
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-713
```


Alignment Scores:

Pred. No.: 4,14e-28 Length: 2124
Score: 301.00 Matches: 50
Percent Similarity: 80.52% Conservatives: 12
Best Local Similarity: 64.94% Mismatches: 15
Query Match: 68.10% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-713 (1-2124)

QY 1 CysLysAspCysGlyLysAlaPheLeGlnLysSerAsnLeuLeAArgHisGlnArgThr 20
Db 256 TGTGCCGATTGTGGAAAAGTTCTTTCAGAGTTCTAACCTCATTTCAGCATCGCGGATC 315

QY 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 316 CACCGGGGAAAGCCCTATAGTCGACGAGTGTGGAGAGAGATTTAAACAGAGCTCC 375

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCys 60
Db 376 AACCTTATCCAGCACACAGAGATTTCACACCGGGGAGAGCCCTATTCTGCGATGATGT 435

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 436 GCGCGGTGCTTTAGCCAGAGTTCACACCTTATTCAGCACAGAAACCCAC 486

RESULT 5

US-11-136-527-4391
; Sequence 4391, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4391
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4391

Alignment Scores:
Pred. No.: 6.65e-28 Length: 1940
Score: 299.00 Matches: 49
Percent Similarity: 83.56% Conservatives: 12
Best Local Similarity: 67.12% Mismatches: 12
Query Match: 67.65% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-295 (1-1940)

QY 3 AspCysGlyLysAlaPheLeGlnLysSerAsnLeuLeAArgHisGlnArgThrHisThr 22
Db 1445 GARTGTGSAARTCTTTACGCAAACTCAACCCCTTAGTCAACACAGAAATTCATCT 1504

QY 23 GlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerAsnLeu 42
Db 1505 GGAGAAAACCTTACAAATGTGAACAGTGTGAAAGGCYTTTACCAATGTTCAGCCTT 1564

QY 43 ThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCysGlyLys 62
Db 1565 AGAAAACACACAGAGAAATTCACACTGGAGAGAAACCTCAAAATGTGAAGAATGTGGCAGA 1624

QY 63 AlaPheGlyCysProSerAsnLeuArgArgHisGlyArg 75
Db 1625 GCCTTTAACTGCGTTCATCTTTTACTAAACACACAGA 1663

Alignment Scores:

Pred. No.: 1.49e-28 Length: 600
Score: 299.00 Matches: 49
Percent Similarity: 83.56% Conservatives: 12
Best Local Similarity: 67.12% Mismatches: 12
Query Match: 67.65% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-4391 (1-600)

QY 3 AspCysGlyLysAlaPheLeGlnLysSerAsnLeuLeAArgHisGlnArgThrHisThr 22
Db 105 GARTGTGSAARTCTTTACGCAAACTCAACCCCTTAGTCAACACAGAAATTCATCT 164

QY 23 GlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerAsnLeu 42
Db 165 GGAGAAAACCTTACAAATGTGAACAGTGTGAAAGGCYTTTACCAATGTTCAGCCTT 224

QY 43 ThrLysHisLysLysIleHisThrGlyLysProTyrLysCysLysGlnCysGlyLys 62
Db 225 AGAAAACACACAGAGAAATTCACACTGGAGAGAAACCTCAAAATGTGAAGAATGTGGCAGA 284

QY 63 AlaPheGlyCysProSerAsnLeuArgArgHisGlyArg 75
Db 285 GCCTTTAACTGCGTTCATCTTTTACTAAACACACAGA 323

```
Best Local Similarity: 67.53% Mismatches: 17
Query Match: 67.42% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-6532 (1-600)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 19 TGTAAAGGGTGTGGAGGCGCTTATTGAGAAATCAGCCTCATCAGACACAGGAGGC 78
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 79 CATACTGGAGAGAAACCGTACACATGTAAGGAATGTGGAAAGCCTTCAGCGGCAATCA 138
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 139 AACCTGACTGAGCATGAGAAATTCATATTGGAGAGAAACCCCTATAAATGTAACGAGTGT 198
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 199 GGAACAATTTTCAGGCAGACAGTACCTCATCAAAATCATCAACATTCAC 249

RESULT 8
US-11-136-527-2436
; Sequence 2436, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; NUMBER OF SEQ ID NOS: 2005-05-26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2436
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2436

Alignment Scores:
Pred. No.: 6,86e-28 Length: 1578
Score: 298.00 Matches: 52
Percent Similarity: 77.92% Conservative: 8
Best Local Similarity: 67.53% Mismatches: 17
Query Match: 67.42% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-2436 (1-1578)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 997 TGTAAAGGGTGTGGAGGCGCTTATTGAGAAATCAGCCTCATCAGACACAGGAGGC 1056
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1057 CATACTGGAGAGAAACCGTACACATGTAAGGAATGTGGAAAGCCTTCAGCGGCAATCA 1116
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1117 AACCTGACTGAGCATGAGAAATTCATATTGGAGAGAAACCCCTATAAATGTAACGAGTGT 1176
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1177 GGAACAATTTTCAGGCAGACAGTACCTCATCAAAATCATCAACATTCAC 1227

RESULT 9
US-10-185-54987/c
; Sequence 54987, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54987
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Bovine 19866880605923
US-10-750-623-54987/c
; Sequence 54987, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54987
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Bovine 19866880605923
US-10-750-623-54987/c

Alignment Scores:
Pred. No.: 6.78e-28 Length: 984
Score: 296.00 Matches: 51
Percent Similarity: 79.22% Conservative: 10
Best Local Similarity: 66.23% Mismatches: 16
Query Match: 66.97% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-185-54987 (1-984)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 609 TGCACCTAAGTGTGGAAAGGCTTTTCAGTAAATCCAGCTTGTGAGCATCAGAGAACT 550
Qy 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 549 CATACAGGTGAAACCCCTACGAGTGGGTGAATGTGGAAAGCTTTTCAGTGAATAATTA 490
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 489 AGCCTCACTAATCACCAGAGAAATTCATACAGGAGAAACCGTATGTGTGAGTGAAGTGT 430
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 429 GGGAAAGCCTTTTGTGCAAGAGTCACTCATCATCATCAGAGGACACAC 379

RESULT 10
US-10-750-623-54987/c
; Sequence 54987, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54987
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Bovine 19866880605923
US-10-750-623-54987
```

Alignment Scores:
Pred. No.: 6.78e-28 Length: 984
Score: 296.00 Matches: 51
Percent Similarity: 79.22% Conservative: 10
Best Local Similarity: 66.23% Mismatches: 16
Query Match: 66.97% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-54987 (1-984)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAtrGHisGlnArgThr 20
Db 609 TGCCTAAGTGTGGGAAGCCTTCAGTAGGAATCCAGCTTGTCAGGCATCAGAGAAT 550

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 549 CATACAGTGAACCAACCTACAGTGGGTGAATGTGGGAAGCTTCAGTGGAATAA 490

QY 41 AsnLeuThrLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
Db 489 AGCCTCACTAATACAGAGAAATTCATACAGAGAAACCGTATGTGTGAGTGAGT 430

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyAtrGThrHis 77
Db 429 GGGAAAGCCTTTGTCAGAACTCACATCTCATATCACAATCAGAGACACAC 379

RESULT 11
US-10-750-185-44166/c
; Sequence 44166, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44166
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Bovine 19866880546157
US-10-750-185-44166

Alignment Scores:
Pred. No.: 8.01e-27 Length: 4305
Score: 294.00 Matches: 50
Percent Similarity: 80.52% Conservative: 12
Best Local Similarity: 64.94% Mismatches: 15
Query Match: 66.52% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-44166 (1-4305)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAtrGHisGlnArgThr 20
Db 3428 TGTAAAGAAATGTAGGAAGCCTTCAGCCAAATGTATCTTATTCATCAATCAAGAAAT 3369

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 3368 CATACCTGGAAGAAACCATATAATGTAGGAGTGTAGAAAAGCCTTCAGCCAGCTGCA 3309

QY 41 AsnLeuThrLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
Db 3308 CACCTTGTCTCAGCATCAGAGAATTCATCTGGGGAGAAAGCCTTATAAAATGTAAGGAAT 3249

RESULT 13
US-11-136-527-6854
; Sequence 6854, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2

Alignment Scores:
Pred. No.: 6.78e-28 Length: 984
Score: 296.00 Matches: 51
Percent Similarity: 79.22% Conservative: 10
Best Local Similarity: 66.23% Mismatches: 16
Query Match: 66.97% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-54987 (1-984)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAtrGHisGlnArgThr 20
Db 609 TGCCTAAGTGTGGGAAGCCTTCAGTAGGAATCCAGCTTGTCAGGCATCAGAGAAT 550

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 549 CATACAGTGAACCAACCTACGAGTGGGTGAATGTGGGAAGCTTCAGTGGAATAA 490

QY 41 AsnLeuThrLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
Db 489 AGCCTCACTAATACAGAGAAATTCATACAGAGAAACCGTATGTGTGAGTGAGTGT 430

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis 77
Db 429 GGGAAAGCCTTTGTCAGAACTCACATCTCATATCATACATCAGAGACACAC 379

RESULT 11
US-10-750-185-44166/c
; Sequence 44166, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44166
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Bovine 19866880546157
US-10-750-185-44166

Alignment Scores:
Pred. No.: 8.01e-27 Length: 4305
Score: 294.00 Matches: 50
Percent Similarity: 80.52% Conservative: 12
Best Local Similarity: 64.94% Mismatches: 15
Query Match: 66.52% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-623-44166 (1-4305)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAtrGHisGlnArgThr 20
Db 3428 TGTAAAGAAATGTAGGAAGCCTTCAGCCAAATGTATCTTATTCATCAATCAAGAAAT 3369

QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 3368 CATACTGGAGAAACCATATAATGTAAAGAGTGTAGAAAAGCCTTCAGCCAGCTGCA 3309

QY 41 AsnLeuThrLysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
Db 3308 CACCTTGTCTCAGCATCAGAGAATTCATCTGGGGAGAAAGCCTTATAAAATGTAAGGAATGT 3249

RESULT 13
US-11-136-527-6854
; Sequence 6854, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2

```
; SEQ ID NO 6854
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-6854

Alignment Scores:
Pred. No.: 2,57e-27 Length: 1400
Score: 293.00 Matches: 51
Percent Similarity: 80.52% Conservative: 11
Best Local Similarity: 66.23% Mismatches: 15
Query Match: 66.23% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-6854 (1-1400)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1047 TGTGAAGAATGTGGAAAGCCCTTGGCTGTAAATCAAACTTTATCGGCATCAGAGGATC 1106
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1107 CATACCTGGTGAACCCCTATCAGTGTAAATCAGTGTGGCAAGCCCTTCAGCCAGTATTCA 1166
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1167 TTCCTAACGGACACGACGCGAATCCATCTCGAGAGAGCTCTACAGTGTATGGAGTGT 1226
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1227 GGGAAAGCCTATAGTTACAGGTCAAACTCTGTAGACACAAAGAAAGTCCAC 1277

RESULT 14
US-11-136-527-2758
; Sequence 2758, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2758
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2758

Alignment Scores:
Pred. No.: 3,25e-27 Length: 1683
Score: 293.00 Matches: 51
Percent Similarity: 80.52% Conservative: 11
Best Local Similarity: 66.23% Mismatches: 15
Query Match: 66.23% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-2758 (1-1683)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1330 TGTGAAGAATGTGGAAAGCCCTTGGCTGTAAATCAAACTTTATCGGCATCAGAGGATC 1389
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1390 CATACCTGGTGAACCCCTATCAGTGTAAATCAGTGTGGCAAGCCCTTCAGCCAGTATTCA 1449
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1449 TGTGAAGAATGTGGAAAGCCCTTGGCTGTAAATCAAACTTTATCGGCATCAGAGGATC 1506

; SEQ ID NO 6854
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-6854

Alignment Scores:
Pred. No.: 2,57e-27 Length: 1400
Score: 293.00 Matches: 51
Percent Similarity: 80.52% Conservative: 11
Best Local Similarity: 66.23% Mismatches: 15
Query Match: 66.23% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-11-136-527-6854 (1-1400)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 1047 TGTGAAGAATGTGGAAAGCCCTTGGCTGTAAATCAAACTTTATCGGCATCAGAGGATC 1106
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 1107 CATACCTGGTGAACCCCTATCAGTGTAAATCAGTGTGGCAAGCCCTTCAGCCAGTATTCA 1166
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 1167 TTCCTAACGGACACGACGCGAATCCATCTCGAGAGAGCTCTACAGTGTATGGAGTGT 1226
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1227 GGGAAAGCCTATAGTTACAGGTCAAACTCTGTAGACACAAAGAAAGTCCAC 1277

RESULT 15
US-10-750-185-41858
; Sequence 41858, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41858
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Bovine 19866881116464
US-10-750-185-41858

Alignment Scores:
Pred. No.: 1,05e-27 Length: 438
Score: 291.00 Matches: 51
Percent Similarity: 72.73% Conservative: 5
Best Local Similarity: 66.23% Mismatches: 21
Query Match: 65.84% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x US-10-750-185-41858 (1-438)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 38 TGTGATGACTGTGGCAAGCCCTTCGTTTAAAGTCATCTCTTTACTCATCAGACAGTT 97
QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 98 CATACCTGGAGAGAAACCTTACAAATGTGATGAGTGTGGCAAGCCCTTCGTCGCAAGTCA 157
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 158 ACCCTTTTATCTCATCAGACAATACATACCTGGAGAGAAACCTTACAAATGTGATGAGTGT 217
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 218 GCGAAGGTCTTTACTCAGCTCAAAATCTCAGGAGACATCAGAAAAATTCAT 268

Search completed: January 12, 2006, 03:46:18
Job time : 228 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 12, 2006, 00:39:24 ; Search time 3025 Seconds

(without alignments)
1190.944 Million cell updates/sec

Title: US-10-669-861-2_COPY_33_109

Perfect score: 442

Sequence: 1 CKDCGKATQKSLRHORT.....KQCGKAFGCPNLRHGRTH 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10669861/runat_09012006_144453_20382/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOCPCL=0 -LOCPXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10669861@CNC 1 1 5315 @runat_09012006_144453_20382 -NCPUS=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	329	74.4	458	1 AJ673279	AW236498
2	327	74.0	518	1 AW236498	CR849534
3	326	73.8	889	7 CR849534	CZ921322
4	326	73.8	922	10 CZ921322	BI654450
5	316	71.5	923	3 BI654450	BG245794
6	316	71.5	990	2 BG245794	BG262051
7	316	71.5	1015	2 BG262051	

ALIGNMENTS

RESULT 1

AJ673279

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 458)

Anderson, S.I., Finlayson, H.A. and Archibald, A.L.

Development of cDNA and EST resources for studying reproduction and

embryo development in pigs and cattle

Unpublished (2004)

Contact: Anderson SI

Genomics and Bioinformatics

Roslin Institute

Roslin, Midlothian, EH25 9PS, UNITED KINGDOM

Single pass sequencing. Bases called and trimmed with phred

v0.020425.c. Vector identified by cross match with the -minscore 20

and -minmatch 12 options. Vector: pBluescriptII(SK+) R. Site 1:

EcORI R. Site 2: NotI 5' Seq primer M13F Description: Normalised

library constructed from Bovine Uterus tissue. Clones available

from UK Centre for Functional Genomics in Farm Animals, Roslin

Institute, Roslin, Midlothian, UK, EH25 9PS, www.arkgenomics.org.

Location/Qualifiers

1..458

/organism="Bos taurus"

CFI145979	UI-HF-CB0	536	6	CFI145979	71.3	315	71.3
CB409262	NISC.nc04	632	6	CB409262	71.3	315	71.3
AUI23664	AUI23664	827	1	AUI23664	71.3	315	71.3
BQ221286	AGENCOURT	875	3	BQ221286	71.3	315	71.3
BM904772	AGENCOURT	1066	3	BM904772	71.3	315	71.3
DQ030160	Homo sapi	1362	11	DQ030160	71.3	315	71.3
BX483171	DKFZp686B	441	5	BX483171	71.0	314	71.0
DQ045809	Pan trogl	584	11	DQ045809	71.0	314	71.0
BQ306591	MS0-BF300	592	3	BQ306591	71.0	314	71.0
AQ356926	CTBTBI-E1	597	9	AQ356926	71.0	314	71.0
AW970087	EST382168	675	1	AW970087	71.0	314	71.0
BM973066	UI-CF-ECL	691	3	BM973066	71.0	314	71.0
CA509386	UI-R-FSO-	724	6	CA509386	71.0	314	71.0
CB325611	UI-R-DZ0-	765	6	CB325611	71.0	314	71.0
BU569663	AGENCOURT	780	5	BU569663	71.0	314	71.0
CD623771	56084457J	807	6	CD623771	71.0	314	71.0
BG482615	602502762	830	2	BG482615	71.0	314	71.0
CJ011854	CJ011854	833	7	CJ011854	71.0	314	71.0
CJ021276	CJ021276	843	7	CJ021276	71.0	314	71.0
CJ038720	CJ038720	847	7	CJ038720	71.0	314	71.0
BG619358	602619493	855	2	BG619358	71.0	314	71.0
CR981016	CR981016	865	7	CR981016	71.0	314	71.0
BP438052	BP438052	874	3	BP438052	71.0	314	71.0
CJ012987	CJ012987	875	7	CJ012987	71.0	314	71.0
DQ045808	Homo sapi	879	11	DQ045808	71.0	314	71.0
BP463090	BP463090	884	3	BP463090	71.0	314	71.0
BP439903	BP439903	887	3	BP439903	71.0	314	71.0
BP433982	BP433982	890	3	BP433982	71.0	314	71.0
BM541411	AGENCOURT	1102	3	BM541411	71.0	314	71.0
DQ032737	Homo sapi	1580	11	DQ032737	71.0	314	71.0
DQ036016	Homo sapi	1836	11	DQ036016	71.0	314	71.0
BC005154	Homo sapi	2197	4	BC005154	71.0	314	71.0
CR859112	Pongo pyg	3126	4	CR859112	71.0	314	71.0
CR857109	Pongo pyg	3239	4	CR857109	71.0	314	71.0
CB809285	AMGNNUC:N	382	6	CB809285	70.8	313	70.8
BB834711	BB834711	461	2	BB834711	70.8	313	70.8
AI788680	uk52f01.x	527	1	AI788680	70.8	313	70.8
AZ998278	2M0285G08	540	9	AZ998278	70.8	313	70.8

```
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="KN224-011_117"
/tissue_type="uterus"
/clone_lib="KN224"
/note="Vector: pBluescriptII(SK+); Site 1: EcoRI; Site 2:
NotI; Single pass sequencing. Normalised library
constructed from Bovine Uterus tissue."

ORIGIN

Alignment Scores:
Pred. No.: 4.46e-28 Length: 458
Score: 329.00 Matches: 56
Percent Similarity: 85.71% Conservative: 10
Best Local Similarity: 72.73% Mismatches: 11
Query Match: 74.43% Indels: 0
DB: 1 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AJ673279 (1-458)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 27 TGCACAGATTGTCACCAAGCCCTTCAGTCACCGCTTATCCGTCATCAGAGAAC 86
Qy 21 HisThrGlyLysGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 87 CACACTGGAGAGAACCTTACAGTGTAGGACTGTGGAAAGCCCTTCAGCAGAGCTCA 146
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysGlyLysProTyrLysCysLysGlnCys 60
Db 147 TCTCTTCAACAGCATCAGAAACTCACACTGGAGAAACCCCTATAAGTGTAAAGGAATGT 206
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 207 GGAAGGGCTTTCAGCCAGCAGTTCATCCCTTTCTCAACATCAGAAACTCAT 257

RESULT 2
AW236498/c AW236498 518 bp mRNA linear EST 13-DEC-1999
LOCUS xm47a05.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2687312 3
DEFINITION similar to SW:Z136 HUMAN P52737 ZINC FINGER PROTEIN 136. ;contains
Alu repetitive element;; mRNA sequence.
ACCESSION AW236498
VERSION AW236498.1 GI:6568887
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
1 (bases 1 to 518)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 400.
Location/Qualifiers
1..518
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
/db_xref="taxon:9913"
/clone="KN224-011_117"
/tissue_type="uterus"
/clone_lib="KN224"
/note="Vector: pBluescriptII(SK+); Site 1: EcoRI; Site 2:
NotI; Single pass sequencing. Normalised library
constructed from Bovine Uterus tissue."

ORIGIN

Alignment Scores:
Pred. No.: 8.87e-28 Length: 518
Score: 327.00 Matches: 56
Percent Similarity: 81.82% Conservative: 7
Best Local Similarity: 72.73% Mismatches: 14
Query Match: 73.98% Indels: 0
DB: 1 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x AW236498 (1-518)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr 20
Db 449 TGTAAAGCAGTGTGGGAAGCCCTTTGGATCTCCCTCACACCTTCAATGCATGGAGGACT 390
Qy 21 HisThrGlyLysGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 389 CACACTGGAGAGAACCTTATGAATGTAAAGCAGTGTGGGAAGCTTTTGGATGTGCTCG 330
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysGlyLysProTyrLysCysLysGlnCys 60
Db 329 CGACTTCAATGCATGGAAGGACTCACACTGGAGAGAACCGTATAATGTAAGCAATGT 270
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 269 GGAAGGCTTTTGGATGTCCCTCAACCTTCGAAGGCAATGGAGGACTCAC 219

RESULT 3
CR849534 889 bp mRNA linear EST 23-MAR-2005
LOCUS CR849534
DEFINITION CR849534 Normalized and Subtracted bovine endometrium tissues
(bcbp) Bos taurus cDNA clone bcbp0003.f.09 5', mRNA sequence.
ACCESSION CR849534
VERSION CR849534.2 GI:61725063
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 889)
Renard J.P., Lewin H.A., Yang J., Hernandez, A., Sandra, O.,
Everts, R.E. and Hue, I.
Endometrium ESTs (bcbp)
Unpublished (2004)
On Oct 21, 2004 this sequence version replaced gi:54564405.
Contact: Renard JP
Biologie du Developpement et Reproduction
INRA
Domaine de Vilvert 78350 Jouy en Josas, FRANCE
Email: renard@jouy.inra.fr
at signasupport@jouy.inra.fr to obtain the chromatogram of this
sequence. AGENAE Resource centre. Francois PIUMI,
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du
Genome (UREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,
FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73
Plate: bcbp0003c row: f column: 09.
Location/Qualifiers
```

```

source
1. 889
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="bcbp0003.f.09"
/tissue_type="caruncular and inter-caruncular regions from
embryo mortality, day 64 pregnant cows, cows displaying
embryo mortality, day 20 cyclic cows"
/clone_lib="Normalized and subtracted bovine endometrium
tissues (bcbp)"
/notes="Vector: pBluescript II SK(+); Funding for
construction and subtraction of the cDNA library was
provided to J. P. Renard (BDR, INRA). Funding for EST
sequencing was provided by a Genanimal contract from
FNS-Ministere de la Recherche (INRA, Num: A02085). The
cDNA library was constructed, normalized and subtracted by
Dr. A. Hernandez (W.M. Keck Center, University of Illinois
Urbana-Champaign) and Dr O. Sandra (BDR, INRA) as
described in Bonaldo et al.(1996), Genome Research 6,
791-806. The BCBP library is a normalized library of
endometrium (caruncular and inter-caruncular regions)
subtracted with previously sequenced clones from bovine
placenta (BP), bovine embryo (BE) and bovine spleen (BS)
libraries produced in the laboratory of Dr. H.A. Lewin
(University of Illinois Urbana-Champaign). The double
stranded cDNA was size selected (more than 450 bp). Size
selected cDNA was ligated to EcoRI adaptors, digested with
NotI then directionally cloned into EcoRI-NotI digested
pBluescript II SK(+) phagemid vector. Insert size of the
bcbp was between 600-3500 bp."

ORIGIN
Alignment Scores:
Pred. No.: 2,278-27 Length: 889
Score: 326.00 Matches: 55
Percent Similarity: 84.42% Conservative: 10
Best Local Similarity: 71.43% Mismatches: 12
Query Match: 73.76% Indels: 0
DB: 7 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x CR849534 (1-889)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
Db 66 TGTAAAGATGTGGAAAGCCCTTTATTCAGAGCTCAGACTTATTCACATCAGAGATC 125
QY 21 HisThrGlyLysProTyrlsCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 126 CATACAGGTGAAAAACCATATGATGATGAGTGTGGAAGGCTTTTAAATAAGGCTCA 185
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrlsCysGlnCys 60
Db 186 AACCTTACTGTCATCAAGAATTCACACTGGTGAGAACCTTATGACTGTAAAGATGT 245
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuAargHisGlyArgThrHis 77
Db 246 GGAAGTCTCTTGGTAGTCGCTCGACCTCATCCGCCATGAAGGAATCAT 296

RESULT 4
C2921322
LOCUS
DEFINITION
1098415575295 CH243 Ovis aries genomic clone CH243-350C22, genomic
survey sequence.
ACCESSION
C2921322
VERSION
C2921322.1 GI:71943436
KEYWORDS
GSS.
SOURCE
Ovis aries (sheep)
ORGANISM
Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovidae; Bovidae; Caprinae; Ovis.
REFERENCE
1 (bases 1 to 922)
AUTHORS
Kirkness,E.

```

```

TITLE
JOURNAL
COMMENT
Ovine BAC Ends
Unpublished (2005)
Other_GSSs: 1098415567615
Contact: Ewen Kirkness
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Email: ekirknes@tigr.org
Sequences generated at the J. Craig Venter Institute Joint
Technology Center (JCVI/JTC; http://www.venterlinstitute.org/)
Insert Length: 184000 Std Error: 0.00 row: C column: 22
Seq primer: FORWARD
Class: BAC ends
High quality sequence start: 78
High quality sequence stop: 892.
Location/Qualifiers
1. 922
/organism="Ovis aries"
/mol_type="genomic DNA"
/db_xref="taxon:9940"
/clone="CH243-350C22"
/clone_lib="CH243"
/notes="Vector: PTARBAC2.1"

ORIGIN
Alignment Scores:
Pred. No.: 2,378-27 Length: 922
Score: 326.00 Matches: 55
Percent Similarity: 84.42% Conservative: 10
Best Local Similarity: 71.43% Mismatches: 12
Query Match: 73.76% Indels: 0
DB: 10 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x C2921322 (1-922)
QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
Db 302 TGTAAAGATGTGGAAAGCCCTTTATTCAGAGCTCAGACTTATTCACATCAGAGATC 361
QY 21 HisThrGlyLysProTyrlsCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 362 CATACAGGTGAAAAACCATATGATGATGAGTGTGGAAGGCTTTTAAATAAGGCTCA 421
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrlsCysGlnCys 60
Db 422 AACCTTACTGTCATCAAGAATTCACACTGGTGAGAACCTTATGACTGTAAAGATGT 481
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuAargHisGlyArgThrHis 77
Db 482 GGAAGTCTCTTGGTAGTCGCTCGACCTCATCCGCCATGAAGGAATCAT 532

RESULT 5
BI654450
LOCUS
DEFINITION
603281023F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5325509 5',
mRNA sequence.
ACCESSION
BI654450
VERSION
BI654450.1 GI:15568686
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 923)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11825 row: n column: 06
High quality sequence start: 699.

FEATURES

```

1: 1:223
Source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:5325509"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Salt;
Site 2: Not1; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

```

RESULT 6
BC245794

[illegible]

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAW10331 row: i column: 22
High quality sequence stop: 701.
Location/Qualifiers
1..990
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4487133"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Maml"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:					
Pred. No.:	3,76e-26	Length:	990		
Score:	316.00	Matches:	53		
Percent Similarity:	81.82%	Conservative:	10		
Best Local Similarity:	68.83%	Mismatches:	14		
Query Match:	71.49%	Indels:	0		
DB:	2	Gaps:	0		
US-10-669-861-2_COPY_33_109 (1-77) x BG245794 (1-990)					
QY	1	CysLysAAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleArgHisGlnArgThr	20		
Db	114	TGTGCCAGTGTGATGAAGGCATTTCACAAACAGTCATCTCTTAACACATAAAAGAAC	173		
QY	21	HisThrGlyGlutysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer	40		
Db	174	CATACTGGAGAGAACCCCTATGAATGTGACCAATGTGGCAAGCCTTTGCAAGTAATAGT	233		
QY	41	AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60		
Db	234	AATCTGCAAGTGCATAAAANAATGCACACTGGAGAGAACCCCTATGAATGTAAAGCATGT	293		
QY	61	GlyLysAlaPheGlyCysProSerAsnLeuArgHisGlyArgThrHis	77		
Db	294	GGTAAGCCCTTTGGTATCAAAGTGTCCTGCAAAAGCATAAACGAACACAT	344		

RESULT 7

RG262051	1015 bp	mRNA	linear	EST 13-FEB-2001		
LOCUS	602373885F1	NIH_MGC_94	Mus musculus	cdna clone IMAGE:4481397 5',		
DEFINITION	mRNA sequence.					
ACCESSION	RG262051					
VERSION	RG262051.1	GI:12771867				
KEYWORDS	EST.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;					
	Sciurognathi; Muridea; Muridae; Murinae; Mus.					
	1 (bases 1 to 1015)					
	NIH-MGC http://mgc.nci.nih.gov/ .					
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)					
TITLE	Unpublished (1999)					
JOURNAL						
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: The Cepko Laboratory CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMML) CDNA Sequencing by: Incyte Genomics, Inc.					

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10316 row: j column: 22
High quality sequence start: 3
High quality sequence stop: 686.
Location/Qualifiers
1. .1015

FEATURES source

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4481397"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 3 88e-26 Length: 1015
Score: 316.00 Matches: 53
Percent Similarity: 81.82% Conservative: 10
Best Local Similarity: 68.83% Mismatches: 14
Query Match: 71.49% Indels: 0
DB: 2 Gaps: 0

US-10-669-861-2_COPY33_109 (1-77) x BG262051 (1-1015)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 174 TGTGACCATGTGATGAAGCATTTGGCAAAACAGTCATCTCTTAACATAAAGAACT 233
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 234 CATACTGGAGAGAAACCCCTATGAATGTGAGCAATGTGGCAAGCCCTTTCAGATTAATAGT 293
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 294 AATCTGCAAGTCATAAAAAATGACACTGGAGAGAAACCCCTATGAATGTGAAGCAGTGT 353
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 354 GGTAAAGCCTTTGGGTATCAAAAGTGTCTGCAAAAGCATAAAGACACAT 404

RESULT 8
CF145979 536 bp mRNA linear EST 06-AUG-2003
LOCUS UI-HF-CB0-asj-f-01-0-UI.r1 NIH_MGC_210 Homo sapiens cDNA clone
DEFINITION IMAGE:30568668 5', mRNA sequence.
ACCESSION CF145979
VERSION CF145979.1 GI:33261423
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 536)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: bento-soares@uiowa.edu
Tissue Procurement: Tim Ratliff
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.

FEATURES source

Location/Qualifiers
1. .536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30568668"
/tissue_type="CNCAP(3)T-225 cell line"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH MGC 210"
/note="Organ: Prostate; Vector: pT7T3 Pac; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pT7T3 Pac vector. The library tag sequence located between the Not I site and the polyA tail is CCCAC. Tissue was provided by Tim Ratliff."

ORIGIN

Alignment Scores:
Pred. No.: 2 29e-26 Length: 536
Score: 315.00 Matches: 56
Percent Similarity: 80.52% Conservative: 6
Best Local Similarity: 72.73% Mismatches: 15
Query Match: 71.27% Indels: 0
DB: 6 Gaps: 0

US-10-669-861-2_COPY33_109 (1-77) x CF145979 (1-536)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAArgHisGlnArgThr 20
Db 90 TGTAAAGAATGTGGCAAGCATTTATGTCTTCAACCCCTAAGTACATAGAGAGATA 149
QY 21 HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
Db 150 CACACTGGAGAGAACCCCTACAAATGTGAAGATGTGGCAAGCCTTTAGCCGTCCTCA 209
QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
Db 210 ACCCTTACTAGCATTAAGCAATTCATCTGGAGAGAAACCCCTACAAATGTAAAGATGT 269
QY 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 270 GGCAAGCCTTTAAGCACTCTCTGAGCAATGTGCAAAAGCATTAATAATACAT 320

RESULT 9 CF409262/c

LOCUS CB409262 632 bp mRNA linear EST 24-MAR-2003
DEFINITION NISC_nc04a08.x1 COGENE 6E MAX Homo sapiens cDNA clone IMAGE:5796015 3', mRNA sequence.

ACCESSION CB409262
VERSION CB409262.1 GI:29166131
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 632)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

cdna Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM13210 row: o column: 01
 High quality sequence stop: 390.

FEATURES
 source
 Location/Qualifiers
 1..875
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6015192"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 92"
 /notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 4 21e-26 Length: 875
 Score: 315.00 Matches: 56
 Percent Similarity: 80.52% Conservatives: 6
 Best Local Similarity: 72.73% Mismatches: 15
 Query Match: 71.27% Indels: 0
 DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x BQ221286 (1-875)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
 |||||:::|||||
 Db 96 TGTAAAGAAATGTGGCAAGCAATTTATATGCTCTTCAACCCCTAAGACATAAGAGGATA 155

QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 |||||:::|||||
 Db 156 CACACTGGAGAGAGCCCTACAAATGTGAAGATGTGGCAAGCCTTTAGCCGTTCTCTCA 215

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 |||||:::|||||
 Db 216 ACCCTTACTAAGCATTAAGCAATTTACTTGGAGAGAAACCTTACAAATGTAAGAATGT 275

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuAargHisGlyArgThrHis 77
 |||||:::|||||
 Db 276 GGCAAGCTTTTAAGCACTCTCAGCCCTTGCTAAACATAAATAATACAT 326

RESULT 12
 BM904772
 LOCUS
 DEFINITION
 BM904772 1066 bp mRNA linear EST 12-MAR-2002
 AGENCOURT 6699480 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557577
 5', mRNA sequence.

ACCESSION
 BM904772
 EST.
 BM904772.1 GI:19355151
 Homo sapiens (human)
 Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE
 1 (bases 1 to 1066)
 NIH-MGC <http://imgc.ncbi.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC

cdna Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM13210 row: o column: 01
 High quality sequence stop: 390.

FEATURES
 source
 Location/Qualifiers
 1..1066
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5557577"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_72"
 /notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN
 Alignment Scores:
 Pred. No.: 5.39e-26 Length: 1066
 Score: 315.00 Matches: 53
 Percent Similarity: 81.82% Conservatives: 10
 Best Local Similarity: 68.83% Mismatches: 14
 Query Match: 71.27% Indels: 0
 DB: 3 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x BM904772 (1-1066)

QY 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuIleAargHisGlnArgThr 20
 |||||:::|||||
 Db 109 TGTGAAGAATGTGGCAAGCCTTTACATGCTGTGCGGCCTCCATAAACATAGAGAACT 168

QY 21 HistHrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer 40
 |||||:::|||||
 Db 169 CATACTGGAGAGAAACCTTACAAATGTGAAGATGTGGCAAGCGTATACATCTCTCA 228

QY 41 AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys 60
 |||||:::|||||
 Db 229 AATCTAACTGAACATAAGACAACTCATCTGGAGAGAAACCTTACAAATGTAAGAATGT 288

QY 61 GlyLysAlaPheGlyCysProSerAsnLeuAargHisGlyArgThrHis 77
 |||||:::|||||
 Db 289 GGCAAGCTTTTAAGCTGCTCAGACCTTTAATAAACATAAGAAATTCAT 339

RESULT 13
 DQ030160
 LOCUS
 DEFINITION
 DQ030160 1362 bp DNA linear GSS 02-JUN-2000
 Homo sapiens HCL1584 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.

ACCESSION
 DQ030160
 GSS.
 DQ030160.1 GI:66881364
 Homo sapiens (human)
 Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE
 1 (bases 1 to 1362)
 Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B.,
 Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
 White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
 A Scan for Positively Selected Genes in the Genomes of Humans and
 Chimpanzees
 (er) PLoS Biol. 3 (6), E170 (2005)

JOURNAL
 PUBMED
 15869325
 REFERENCE
 2 (bases 1 to 1362)
 Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B.,
 Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D.,
 White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
 A Scan for Positively Selected Genes in the Genomes of Humans and
 Chimpanzees
 (er) PLoS Biol. 3 (6), E170 (2005)

White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.

FEATURES Location/Qualifiers

source 1..1362
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
gene <1..>1362
/locus_tag="HC1584"

ORIGIN

Alignment Scores:
Pred. No.: 7,3e-26 Length: 1362
Score: 315.00 Matches: 53
Percent Similarity: 81.82% Conservative: 10
Best Local Similarity: 68.83% Mismatches: 14
Query Match: 71.27% Indels: 0
DB: 11 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x DQ030160 (1-1362)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
Db 1108 TGTGAAGATGTGGCAAGCCCTTTACATGGTCTGCAGGCCCTCCATAAATAGAGAACT 1167
Qy 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 1168 CATACTGAGAGAAACCTTACAAATGTGAAGATGTGCAAGCCGTATACATCTCTCA 1227
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
Db 1228 AATCTAACTGAACATAAGCAACTCATACTGGAGAGAAACCTTACAAATGTGAAGAAATGT 1287
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 1288 GGCAAGCTTTTAACTGGTCTCAGACCTTAAATAACATAAGAGAAATTCAT 1338

RESULT 14
BX483171 441 bp mRNA linear EST 04-SEP-2003
LOCUS DKFP686B22235_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION DKFP686B22235 5', mRNA sequence.
ACCESSION BX483171
VERSION BX483171.1 GI:31943773
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 441)
Fobbo,G., Boecker,H., Mewes,H.W., Weil,B., Amid,C., Oeanger,A.,
EST (Boecker,H., Boecker,H., Mewes,H.W., Weil,B., Amid,C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFP686B22235) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

FEATURES Location/Qualifiers

source 1..441
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP686B22235"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Alignment Scores:
Pred. No.: 2,35e-26 Length: 441
Score: 314.00 Matches: 54
Percent Similarity: 80.52% Conservative: 8
Best Local Similarity: 70.13% Mismatches: 15
Query Match: 71.04% Indels: 0
DB: 5 Gaps: 0

US-10-669-861-2_COPY_33_109 (1-77) x BX483171 (1-441)

Qy 1 CysLysAspCysGlyLysAlaPheIleGlnLysSerAsnLeuLeuAArgHisGlnArgThr 20
Db 124 TGCAGTGAATGTGTAAGCCCTTCATTCAGAGCAATTCCTTATTCGCACACAGAAAGT 183
Qy 21 HisThrGlyLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSer 40
Db 184 CACACTGGAGAAAAACCTATAAATGCAATGATGTGGAAAGGCTTTAATCAGAACACC 243
Qy 41 AsnLeuThrLysHisLysLysIleHisThrGlyLysProTyrLysCysGlnCys 60
Db 244 TGCCTCACTCAGCATATGAGAAATTCATCTGGAGAGAGCCCTATAAATGTAAAGAATGT 303
Qy 61 GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis 77
Db 304 GGGAAAGCCCTTGTCTCATAGCTCATCTCTTACTGAACATCATAGACTCAC 354

RESULT 15
DQ045809 584 bp DNA linear GSS 02-JUN-2005
LOCUS DQ045809
DEFINITION Pan troglodytes ZNF146 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION DQ045809
VERSION DQ045809.1 GI:66897024

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.

REFERENCE 1 (bases 1 to 584)
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civallo,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (er) PLOS Biol. 3 (6), E170 (2005)

PUBMED 15869325

REFERENCE 2 (bases 1 to 584)

Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civallo,D.,
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.

FEATURES Location/Qualifiers

source 1..584

```
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>584
/gene="ZNF146"
/locus_tag="HC11576"
```

Gene

ORIGIN

Alignment Scores:

Pred. No.:	3.33e-26	Length:	584
Score:	314.00	Matches:	54
Percent Similarity:	80.52%	Conservative:	8
Best Local Similarity:	70.13%	Mismatches:	15
Query Match:	71.04%	Indels:	0
DB:	11	Gaps:	0

US-10-669-861-2_COPY_33_109 (1-77) x DQ045809 (1-584)

QY	1	CysLysAspCysGlyLysAlaPheLeGlnLysSerAsnLeuIleArgHisGlnArgThr	20
DB	304	TGTAAGATTGGCGGAAGCTTTTCATTCAAGAGTCAACCTCATCAGACACCCAGAGAACT	363
QY	21	HisThrGlyGluLysProTyrLysCysGluGluCysGlyLysAlaPheThrGlnSerSer	40
DB	364	CACACAGGAGAGAGCCCTTTGTATGTAAGAGTGTGGAACCTTCAGTGGCAATCC	423
QY	41	AsnLeuThrLysHisLysLysIleHisThrGlyGluLysProTyrLysCysLysGlnCys	60
DB	424	AACCTTACTGACATCAGAGAAATCCATATTGGAGAGAGCCCTTTAAATGTAGTGAATGT	483
QY	61	GlyLysAlaPheGlyCysProSerAsnLeuArgArgHisGlyArgThrHis	77
DB	484	GGAACAGCCCTTGGCCAGAGAAGTACCTCATATAAACATCAAAACATTCAC	534

Search completed: January 12, 2006, 03:29:25
Job time : 3029 secs

This Page Blank (uspto)